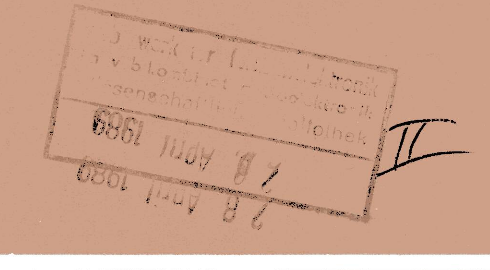
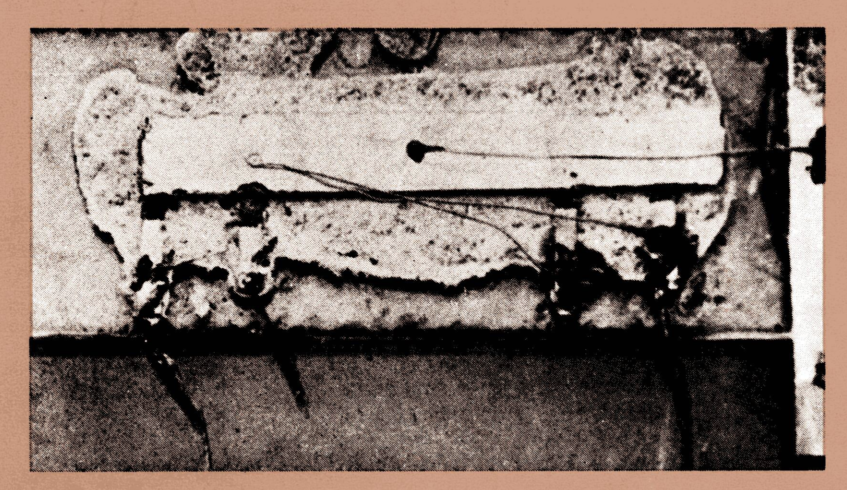
Nachrichtentechnik Elektronik

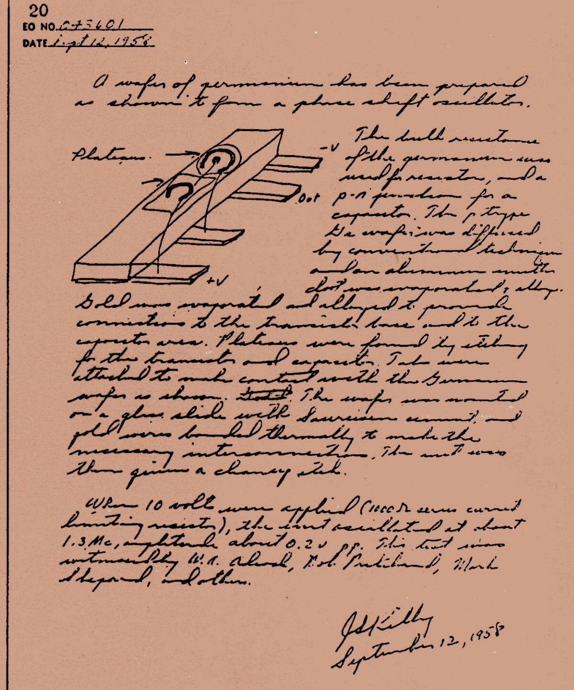
4

1989

VEB VERLAG TECHNIK BERLIN · ISSN 0323-4657







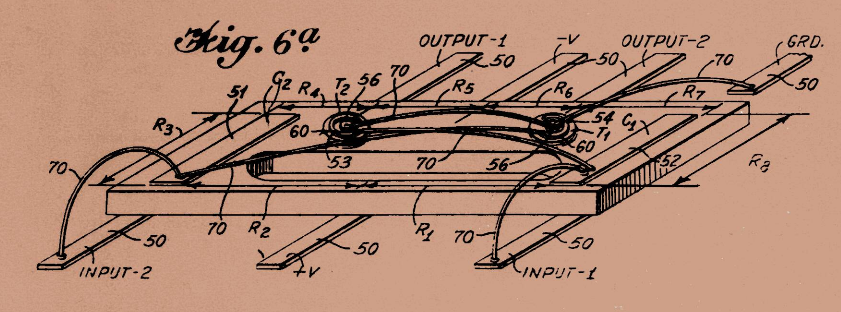
20
EO NO. C+3601
DATE Sept 12, 1958

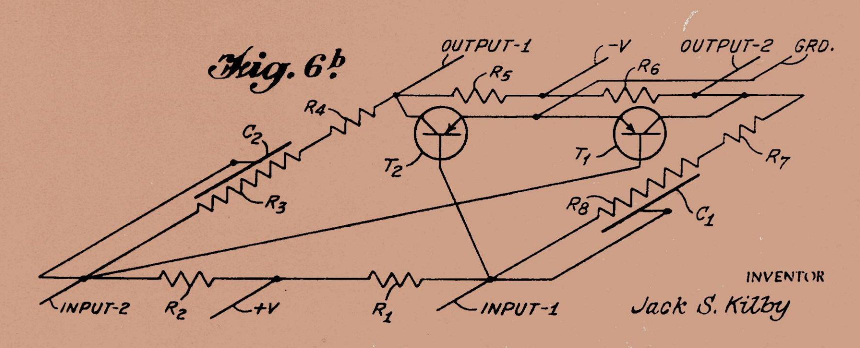
A wafer of germanium has been prepared as shown to form a phase shift oscillator.

The bulk resistance of the germanium was used for resistor, and a p-n junction for a capacitor. The p type Ge wafer was diffused by conventional techniques and an aluminum and gold dot was evaporated, alloy. Gold was evaporated and alloyed to provide connections to the transistor base and to the capacitor area. Plateaus were formed by etching for the transistor and capacitor. Tabs were attached to make contact with the Germanium wafer as shown. The wafer was mounted on a glass slide with Sauereisen cement and gold wires bonded thermally to make the necessary interconnections. The unit was then given a cleaning etch.

When 10 volts were applied (1000 Ω series current limiting resistor), the unit oscillated at about 1.3 Mc, amplitude about 0.2 v pp. This test was witnessed by W.A. Adcock, P.J. Pritchard, Mark Shepard, and others.

J S Kilby
September 12, 1958





30 Jahre integrierte Schaltkreise

СОДЕРЖАНИЕ

CONTENTS

SOMMAIRE

CONTENIDO

4
89

39. Jahrgang

VEB VERLAG TECHNIK
Träger des Ordens
Banner der Arbeit

Nachrichtentechnik Elektronik

Wissenschaftlich-technische Zeitschrift
für die gesamte elektronische Nachrichtentechnik



Herausgeber:
Kammer der Technik
FV Elektrotechnik

Vorschau

Im Heft 5 werden u. a. folgende Beiträge veröffentlicht:

- Applikationserfahrungen mit dem Lasermodul VQ 150 in LWL-Übertragungssystemen PCM 480
- Zusatzdämpfung durch Mikrokrümmungen in kurzen Multimode-Lichtwellenleitern mit Stufenprofil
- Rauschanalyse von Multirate-Systemen. Teil 1 Anwendungen fehlerkorrigierender Codes
- Koaxialkabel für das Frequenzband bis 12,4 GHz und ihre Herstellung in der CSSR

Inhalt

ISSN 0323-4657

Umschau

Titelbild

Vor 30 Jahren erfanden J. Kilby und R. Noyce die integrierte Schaltungstechnik. Auf Seite 122 beginnt aus diesem Anlaß ein technikhistorischer Bericht in 3 Teilen.
(Fotos: Archivaufnahmen)

30 Jahre integrierte Schaltkreise – Entstehung und Entwicklung der Mikroelektronik. Teil 1

H.-P. Bär; A. Kirpal, Ilmenau

Mitteilung aus der Technischen Hochschule Ilmenau

1. Elektronik und Mikroelektronik

Wohl kaum ein Gebiet der Technik hat in den vergangenen 30 Jahren einen solch raschen Aufschwung und entscheidenden Einfluß auf andere Technikentwicklungen genommen wie die Mikroelektronik. Computergesteuerte Maschinen in Verbindung mit Robotern wurden erst durch den Übergang zu immer kleineren Strukturen der Steuer- und Regeleinrichtungen bei gleichzeitiger Erhöhung der Leistungsfähigkeit und Verringerung der Kosten möglich. So wird die Elektronik, vor allem in ihrer bisher höchsten Form, der Mikroelektronik, zur Schlüsseltechnik für die Bewältigung des technischen Kernprozesses der wissenschaftlich-technischen Revolution. Sie greift mit ihren nahezu ungeahnten technischen Möglichkeiten entscheidend in alle Bereiche des gesellschaftlichen Reproduktionsprozesses ein, revolutioniert Produktion und Konsumtion auf der Grundlage neuer Formen und Dimensionen der Informationsgewinnung und Informationsverarbeitung und schafft zugleich Voraussetzungen für die Vervielfachung der geistigen Wesenskräfte des Menschen. Eine ökonomisch sinnvolle und technisch machbare Verarbeitung der zunehmend massenhaft anfallenden Informationen verlangt nach geeigneten technischen Geräten und entsprechenden Verarbeitungsvorschriften. Deshalb sind die Entwicklung der Informatik, der Informationstechnik, der Steuer-, Meß- und Regeltechnik untrennbar mit dem Fortschreiten der Elektronik verbunden, sie stellen heute die entscheidenden Anforderungen an die Entwicklung der Mikroelektronik.
Zur Verdeutlichung des technischen Fortschritts ist es üblich, die Elektronik historisch nach der Entwicklung der bestimmenden Bauelemente und der darauf beruhenden spezifischen Schaltungseigenschaften zu unterteilen:

1. Generation
: durch die Elektronenröhre gekennzeichnet, diskrete Bauelemente, dreidimensionale Verdrahtung der Schaltungen, Trennung von Bauelement und Schaltung

2. Generation
: diskrete Halbleiterbauelemente, Transistor als bestimmendes elektronisches Bauelement, gedruckte Schaltungen, zweidimensionale Verdrahtung der Schaltungen, Trennung von Bauelement und Schaltung

3. Generation
: Mikroelektronik, integrierte Bauelemente, verschiedene Integrationstechniken, technologische Prozesse wie bei den Halbleiterbauelementen der 2. Generation, zweidimensionale Verdrahtung der Schaltungen, beginnende Einheit von Bauelement und Schaltung

4. Generation
: Halbleiterblocktechnik, hochintegrierte Schaltungen, Einheit von Bauelement und Schaltung, Großintegration.

Integrierte Schaltungen auf der Basis der Siliziumchip-Technologie, sind heute zur Standardtechnik der Mikroelektronik geworden. Bestand nach der Erfindung des Transistors und seinem massenhaften Einsatz in der Elektronik, dieser Zustand wurde etwa Ende der 50er Jahre erreicht, die Forderung darin, durch Verringerung der Bauelemente und der Verbindungspunkte die Zuverlässigkeit elektronischer Schaltungen weiter zu erhöhen, lag das weitere Bestreben der Bauelementeentwickler vor allem darin, immer mehr Einzelelemente je Bauelement bzw. je Baugruppe zu verwirklichen. Die Entwicklung führte über verschiedene Miniaturisierungskonzepte 1958 zum monolithisch integrierten Halbleiterschaltkreis. Bei ihm werden im Unterschied zu den vorhergehenden Konzepten alle elektronischen Bauelemente in einem Stück des Halbleitermaterials (Chip) realisiert.

Die Steigerung der
- Integrationsdichte (Elemente je Flächeneinheit) in Verbindung mit einer
- Erhöhung des Integrationsgrades (Elemente je Chip) ermöglichte immer billigere Funktionseinheiten und damit den massenhaften Einsatz monolithischer Halbleiterschaltungen.

Beide Prozesse sind bei weitem noch nicht abgeschlossen. Mitunter spricht man bereits von einer

5. Stufe der Integration:
: ULSI-Technik (ultra large scale integration), die Schaltkreise mit mehr als einer Million Bauelementefunktionen je Chip (sogenannte Megachips) ermöglicht.

Beispiele sind der 1-Mbit-Speicher bzw. der 4-Mbit-Speicher.

Antriebsursache für das Streben nach ständiger Erhöhung der Integrationsdichte und des Integrationsgrades ist die Tatsache, daß die Kosten eines integrierten Schaltkreises im wesentlichen von der benötigten Chipfläche, nicht aber von der Bauelementeanzahl bestimmt werden. Man kann daher durch Verkleinerung der Bauelementestrukturen und Erhöhung des Integrationsgrades die Kosten je elektronischer Funktion immer weiter senken (wobei einschränkend die ständige Erhöhung der Kompliziertheit der technologischen Ausrüstungen dieser Tendenz entgegenwirkt; Bild 1).
Bekanntermaßen sind die Fortschritte bei der weiteren Miniaturisierung und Erhöhung des Integrationsgrades vor allem das Ergebnis verbesserter technologischer Verfahren, die eben die kleineren Strukturen und auch den Einsatz größerer Chipflächen ermöglichen.

Aber auch veränderte Bauelementewirkprinzipien und Schaltungskonzepte unterstützen diese Entwicklungsrichtung. Beispielsweise hat der Ersatz von statischen Schaltungen durch dynamische bei

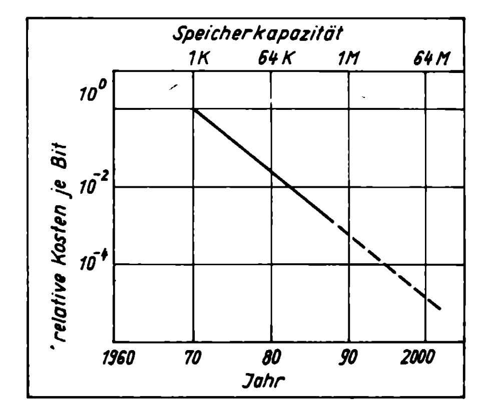


Bild 1
Zusammenhang zwischen den Kosten pro Bauelementefunktion und dem Integrationsgrad [1]

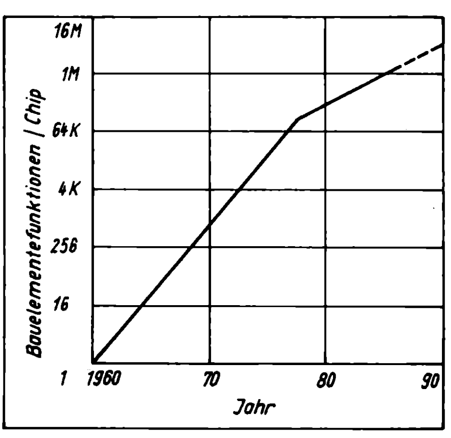


Bild 2
Zeitliche Entwicklung des Integrationsgrades [2] S. 33

Speicherschaltkreisen zu einer wesentlichen Verringerung der Kosten geführt [2]. So zeichnet sich auch immer deutlicher ab, daß die „traditionelle" Unterscheidung zwischen Bauelementeentwickler und Schaltungstechniker weiter verschwindet. Kooperation tritt an ihre Stelle, bis hin zu einer weitestgehenden Verschmelzung im schaltungsorientierten Bauelementeentwurf.

G. Moore fand, daß die zeitliche Entwicklung des Integrationsgrades durch ein Exponentialgesetz beschrieben werden kann, wobei sich von 1960 bis 1978 die Bauelementeanzahl je Chip jährlich verdoppelte.

Ab 1978 war dann eine Verdopplung aller zwei Jahre zu beobachten (Bild 2).

2. Frühere Miniaturisierungs- und Integrationskonzepte in der Elektronik

Der Wunsch nach Miniaturisierung und Zusammenfassung elektronischer Schaltungen ist genau so alt wie die Elektronik selbst. Bereits in der Anfangsphase der Rundfunkempfangstechnik und der Röhrenelektronik gab es Versuche, Elektronenröhren, Widerstände und Kondensatoren zu verkleinern und in Gehäusen, Baugruppen usw. zusammenzufassen, um damit das Volumen der Geräte zu verringern. So vereinigte man bereits in der Röhrentechnik mehrere Röhrensysteme in einem Glaskolben, um Platz und Heizleistung zu sparen. Es gab sogar erfolgreiche Versuche, Widerstände und Kondensatoren mit im Röhrenkolben unterzubringen.

- Am bekanntesten ist wohl die legendäre und unter maßgeblicher Mitwirkung von *M. v. Ardenne* 1926 entstandene Loewe-Dreifachröhre, die einen kompletten Dreiröhrenverstärker in sich vereinigte und in großen Stückzahlen im Ortsempfänger „Loewe OE 333" zum Einsatz kam.

Aber nicht nur die Volumenreduzierung hatten die Erfinder dieser Mehrfachröhre im Sinn, sondern es ging auch darum, durch kürzeste Verbindungen zwischen den im Glaskolben untergebrachten Kopplungsgliedern hohe Verstärkungen auch bei höheren Tonfrequenzen zu erreichen.

M. v. Ardenne bezeichnet diese Dreifachröhre als „ersten integrierten Schaltkreis der Elektronik" [3].

Im Lauf der Entwicklung wurden die Empfängerröhren immer kleiner, es entstanden die Miniaturröhren, ab 1952 sprach man von Subminiaturröhren. Die Zielstellung dieser Miniaturisierungsbestrebungen bestand immer in der Verringerung des Platzbedarfs und der Leistungsaufnahme der Schaltung, sowie in der Verringerung der Anzahl der benötigten Bauelemente je Schaltung.

In dem Maße, wie elektronische Schaltungen immer umfangreicher wurden, vorangetrieben durch die Entwicklung der Rechentechnik, zeigte sich, daß eine Miniaturisierung der Bauelemente allein noch nicht ausreichte – die Zuverlässigkeit der ersten mit Röhren bestückten Rechner wurde durch die Ausfallrate der Röhren und die ungeheure Anzahl von Lötverbindungen begrenzt.

Ende der fünfziger Jahre entwickelte sich der Begriff Miniaturisierung zum Schlagwort der amerikanischen Elektronikindustrie. In der Literatur werden unterschiedliche Gründe dafür genannt, daß diese Bewegung einsetzte. Bereits während des Korea-Krieges 1952 sollen Spezialisten der US-Armee an der Verwirklichung elektronischer Funktionen nach ganz neuen Prinzipien gearbeitet haben. Zielstellung der Arbeiten war, einzelne Moleküle so zusammenzusetzen, daß die Struktur eine elektronische Funktion ausführen kann. Man nannte das Konzept daher Molekularelektronik. Obwohl das Programm 1959 nochmals mit erheblichen Mitteln gestartet wurde, konnten keine greifbaren Erfolge erzielt werden, da man sich mit diesen Vorstellungen damals weit außerhalb der technologischen Realität befand [4].

Eine zweite Antriebsquelle der Miniaturisierungsbestrebungen wird in den Forderungen der Raketentechniker der US-Armee und der amerikanischen Raumfahrtbehörde NASA an die Elektronikindustrie gesehen. Es wurden harte Anforderungen bezüglich Platzbedarf, Leistungsaufnahme und Zuverlässigkeit der Elektronik gestellt [5]. Insbesondere für die Weltraumforschung zu entwickelnden Raketen und Satelliten erforderten

- eine große Zuverlässigkeit der elektronischen Ausrüstung über verhältnismäßig lange Zeiträume.

Gleichzeitig waren

- zunehmende Leistungsfähigkeit und
- Verminderung der Abmessungen und der Masse elektronischer Geräte

zu gewährleisten.

Daraus erwuchs eine besondere Anforderungssituation für die Miniaturisierung an die Elektronik, die in den USA vor allem durch die sogenannte Molekularelektronik bewältigt werden sollte. So sind in den USA durch staatliche Stellen zwischen 1959 und 1965 für Forschung und Entwicklung im Bereich der Integrationstechnik mehr als 100 Millionen Dollar zur Verfügung gestellt worden [6] S. 111.

Neben dem Molekularelektronik-Konzept existieren noch weitere Versuche zur Schaltungsminiaturisierung:

- In der Modultechnik wurden Miniaturwiderstände und -kondensatoren und verkapselte Transistoren auf keramische Platten, auf die zuvor mit Siebdrucktechnik Leiterbahnen aufgebracht worden waren, aufgelötet. Diese Platten wurden übereinandergeschichtet und durch Verlöten von Drähten in Aussparungen an den Plattenkanten zu Schaltungen verbunden und dann komplett verkapselt [7].
- In der Weiterentwicklung zur Mikromodultechnik verringerte man die Plattengröße bis auf wenige Millimeter und stellte Widerstände und Kondensatoren direkt auf den Plättchen durch Siebdruck her. Auf die obere Deckplatte setzte man Miniaturröhrchen (Tinkertoy-Module; USA 1950), später Transistoren in Miniaturgehäusen (Mikromodulplan der US-Armee ab 1959) [8]. Bis zur Einstellung der Produktion 1962 wurden in den USA 5 Millionen derartiger Module produziert und überwiegend für die militärische Elektronik genutzt [9].

Als weitere Varianten gab es

- Dickschicht- und
- Dünnschichttechnik.

Sie wurden seit 1957 entwickelt. Dabei brachte man auf keramische Trägermaterialien Widerstände und Kondensatoren durch Siebdruck bzw. fotolithografische Strukturierung von aufgedampften Schichten auf und lötete nachträglich Halbleiterbauelemente in Miniaturgehäusen auf. Auf eine Übereinanderstapelung mehrerer Platten verzichtete man. Dabei lieferte die Dünnschichttechnik genauere Werte für Widerstände und Kondensatoren und ermöglichte höhere Packungsdichten, war jedoch teurer in der Herstellung als die Dickschichttechnik. Diesem Konzept war zumindest ein Teilerfolg beschieden, in den USA wurden z. B. große Stückzahlen transistorisierter Hörgeräte in Dickschichttechnik hergestellt [9], und auch heute noch werden die Techniken für Spezialanwendungen (z. B. in der Kfz-Elektronik) benutzt.

Wichtige Grundlagen der späteren integrierten Schaltkreistechnologie wurden im Rahmen dieser Entwicklung gelegt:

- Es entstand die Step-and-repeat-Kamera, die es ermöglicht, einen Schaltkreisentwurf immer wieder zu reproduzieren und die für die Maskenherstellung benötigt wird.
- Die fotolithografische Technik wurde weiterentwickelt.

Mitte der 60er Jahre wurden alle diese Konzepte durch die Halbleiterblocktechnik verdrängt, deren Vorteile sehr schnell sichtbar wurden.

Tafel 1 verdeutlicht die Vormachtstellung der monolithischen Integrationstechnik im Vergleich zu anderen Integrationstechniken bereits in den Anfangsjahren der Mikroelektronikentwicklung.

3. Technisch-technologische Voraussetzungen für die Schaltkreisentwicklung

Aus heutiger Sicht kann man feststellen, daß die integrierte Schaltungstechnik eine Erfindung ist, die im Bemühen um die weitere Verbesserung der Transistortechnologie möglich wurde.

Für die massenhafte Anwendung des Transistors als elektronisches Bauelement, wobei er in den meisten Fällen die Elektronenröhre aus ihren „traditionellen Einsatzgebieten" zu verdrängen hatte, galt es, sowohl seine technischen Parameter zu verbessern als auch die Herstellungskosten entscheidend zu senken. Das war in den 50er Jahren nicht so sehr eine Frage anderer Wirkprinzipien, sondern eine Herausforderung an die Transistortechnologie. *J. Auth* ist da-

Tafel 1. Entwicklung der IC-Produktion in Japan [10]

	Gesamt (Mill. Doll.)	Monolith. IC %	Dünnfilm-IC %	Hybrid-IC %	Hybride mikroelektr. Module %
1966	1,6	57,3	41,6	0,2	0,9
1967	8,5	82,3	15,6	0,4	1,7
1968	36,9	77,4	4,1	1,5	16,9
1969	72,5	80,6	3,1	3,6	12,7

her unbedingt zuzustimmen, wenn er in einem Beitrag zum 25jährigen Jubiläum der Erfindung des Transistors schreibt: „Die Entwicklung der Halbleitertechnik vom Erkennen des Transistorprinzips bis zu den heute als Massenprodukt erzeugten Festkörperschaltkreisen ist vor allem eine Entwicklung der Technologie, eine Entwicklung von Bauelementekonstruktionen und dazu passender Herstellungsverfahren" [11]. (In den weiteren Ausführungen wird diese Aussage in bezug auf moderne Integrationstechniken und Bauelementekonzepte relativiert werden.)

- Der entscheidende Durchbruch in der Transistorfertigung gelang mit der Entwicklung des Silizium-Planarprozesses am Ende der 50er Jahre.

Nachdem bereits zu Beginn der 50er Jahre das Diffusionsverfahren zur Erzeugung der pn-Übergänge in den Bell Laboratorien und bei General Electric ausgearbeitet worden war, wurde 1958 durch *J. A. Hoerni* bei Fairchild die Technologie des Si-Planarprozesses bis zur Produktionsreife entwickelt [12]. Erstmals wurde es möglich, durch Ausnutzen der diffusionshemmenden und passivierenden Eigenschaften des Siliziumdioxids und durch die Anwendung fotolithografischer Verfahren der Strukturierung die Bearbeitungsschritte für alle auf der Halbleiterscheibe zu erzeugenden Transistoren gleichzeitig und von einer Seite auszuführen. Durch die Passivierung der an der Oberfläche liegenden pn-Übergänge mit einer Siliziumdioxidabdeckung konnte ein altes Problem der Transistorfertigung zufriedenstellend gelöst werden:

- Gewährleistung einer hohen Zuverlässigkeit der Transistoren und Sicherung der zeitlichen Konstanz ihrer Kennwerte.

Zugleich bot sich die Oxidschicht als ein recht brauchbares Trägermaterial für die aufgedampften Kontaktflächen an.
Die Forschungsarbeiten bei Fairchild wurden vor allem von den acht Wissenschaftlern getragen, mit denen *W. B. Shockley* 1955 die Shockley Semiconductors Laboratories gegründet hatte. Da sich jedoch bei diesem Unternehmen nicht die gewünschten ökonomischen Erfolge einstellten und überdies mit *Shockley* Meinungsverschiedenheiten in bezug auf das Produktionsprogramm und die notwendigen Forschungsarbeiten auftraten, verließ die Mehrzahl der jungen und befähigten Wissenschaftler *Shockleys* Unternehmen bereits nach 2 Jahren. Sie nahmen ihre Tätigkeit bei der zunächst kleinen Firma Fairchild Camera and Instruments auf (später: Fairchild Semiconductor Company). Mit dem Austritt der Wissenschaftler aus *Shockleys* Unternehmen setzte in den USA ein ausgeprägter Wissenstransfer zwischen den Elektronikfirmen ein, getragen von der Absicht, Höchstprofite auf dem sich entwickelnden Halbleitermarkt zu erzielen und immer verbunden mit dem Wechsel hervorragender Fachleute. *Shockley* soll zum Weggang seiner Mitarbeiter später gesagt haben: „Wenigstens eines habe ich getan, ich habe in meinem Urteil recht gehabt: es waren gute Leute, die acht, und meine Leistung war wohl, sie zusammengebracht zu haben" [13] S.170.
Zur Ausarbeitung des Planarprozesses ist noch hinzuzufügen, daß die grundlegenden Versuche zum Oxidmaskenverfahren bereits von *L. J. Derick* und *C. J. Frosh* im Jahre 1955 durchgeführt wurden. Die industrielle Herstellung von Si-Planar-Transistoren wurde 1959 aufgenommen [14] S.86.
Auch andere technologische Entwicklungen der Transistortechnik konnten für die Schaltkreistechnologie übernommen werden:
Genannt seien als Beispiele

- Herstellung von Halbleitermaterialien hoher Perfektion
- Bonden mit Gold- und Aluminiumdrähten
- Vakuumabscheidung von Aluminium- und Goldschichten
- thermische Oxidation des Siliziums
- verschiedene Verkappungstechniken.

Die nachfolgende Aufstellung gibt einen Überblick über wesentliche technologische Entwicklungen des Transistors, vor allem im Hinblick auf die Schaffung von Voraussetzungen für die Schaltkreistechnologie:

1952 industriemäßige Herstellung von Ge-Flächentransistoren nach dem Legierungsverfahren
1954 industriemäßige Herstellung von Si-Flächentransistoren nach dem Legierungsverfahren
1955 grundlegende Untersuchungen zum Oxid-Maskenverfahren
1956 Ausarbeitung des Prinzips des Mesa-Transistors
1958 Ausarbeitung des Si-Planarprozesses für die Herstellung von Bipolartransistoren, industriemäßige Herstellung von diffundierten Si-Flächentransistoren
1959 Entwicklung des Epitaxieverfahrens, industriemäßige Herstellung von Si-Planartransistoren,
Vorschlag für die Realisierung von MIS-Feldeffekttransistoren, Beschreibung des Aufbaus integrierter Schaltkreise (Flip-Flop, Ge) durch *J. Kilby*,
Beschreibung des Aufbaus integrierter Schaltkreise (Anwendung der Planartechnik) durch *R. Noyce*
1960 Beginn der industriemäßigen Herstellung von integrierten Bipolarschaltkreisen (Si-Planartechnik).

1959 ist in den USA etwa ein gleiches Produktionsvolumen für Transistoren und Empfängerröhren (bei Berücksichtigung der Produktion für Ersatzbestückung) erreicht. Die mengenmäßige Gleichheit tritt etwa 1961 ein, obwohl der Durchschnittspreis des Ge-Transistors mit 1,14 Dollar und der des Si-Transistors mit 7,48 Dollar noch deutlich über dem der Elektronenröhre liegt. 1961 werden in den USA etwa 180 Millionen Ge-Transistoren und 13 Millionen Si-Transistoren hergestellt, [15] S.168–174.

4. Die Entwicklung der ersten monolithisch integrierten Halbleiterschaltkreise

Die Entwicklung der integrierten Schaltkreise ist eng mit den Namen *Jack Kilby* und *Robert Noyce* verbunden. Sie haben jedoch nicht gemeinsam, sondern unabhängig voneinander gearbeitet (*Kilby* bei Texas Instruments und *Noyce* bei Fairchild). Um die Urheberrechte haben beide im Nachhinein eine gerichtliche Auseinandersetzung geführt. *Kilby* klagte gegen die Patentanmeldung von *Noyce*. Mit einer Entscheidung des U. S. Court of Costume and Patent Appeals vom November 1969 wurde *R. Noyce* als Erfinder der integrierten Schaltung bestätigt, [6] S.110. Da jedoch beide Forscher unbestreitbare Verdienste um die Erfindung des integrierten Schaltkreises haben, auf die Einzelheiten ihrer Leistungen wird noch ausführlicher eingegangen, wurden *Noyce* und *Kilby* 1966 mit der Ehrenmedaille des amerikanischen Franklin-Instituts ausgezeichnet: *Kilby* für die „Konstruktion und Konzeption des ersten integrierten Schaltkreises", *Noyce* für die „Verfeinerung des ersten monolithischen Schaltkreises für spezielle Verwendung, vor allem in der Industrie", [4] S.45.
Die Grundidee des integrierten Schaltkreises wurde bereits 1952, nur 4 Jahre nach der Erfindung des Transistors, geäußert:
G. W. A. Dummer vom britischen Royal Radar Establishment formulierte auf einer Bauelementetagung in Washington sinngemäß, daß es mit der Erfindung des Transistors nun möglich wäre, elektronische Schaltungen in einem Block ohne Verbindungsdrähte zu realisieren. Der Block sollte aus isolierendem, leitendem, gleichrichtendem und verstärkendem Material bestehen, das durch Heraustrennen bestimmter Gebiete direkt zu verbinden wäre [4] S.45.
Die Frage, warum darauf hin nicht die Engländer den ersten Schaltkreis gebaut haben, ist oft gestellt worden. Auf die vielfältigen Ursachen soll hier nicht näher eingegangen werden; ein Hauptgrund ist wohl, daß die britische Halbleiterindustrie in der Entwicklung weit hinter den USA zurücklag und *Dummers* Idee zudem von staatlichen Stellen nicht in ihrer Bedeutung erkannt wurde.
Ideengeschichtlich interessant ist ebenso die Tatsache, daß bereits *W. B. Shockley* in seiner 1948 eingereichten Patentschrift zum Flächentransistor unter den Anwendungsmöglichkeiten auf „Mehrfachsteuerungen für Hochfrequenzmischstufen" hinweist und entsprechende halbleitende Mehrfachstrukturen angibt („integrierter Schaltkreis") [16].

NaA 134

Doz. Dr. sc. phil. Dr.-Ing. *Alfred Kirpal*, Technische Hochschule Ilmenau, Sektion Marxismus-Leninismus; Dipl.-Ing. *Hans-Peter Bär*, Technische Hochschule Ilmenau, Sektion Physik und Technik Elektronischer Bauelemente, PSF 327, Ilmenau, 6300

Fortsetzung im H. 5

Bemerkungen zum Modellbegriff der Systemtechnik

F. H. Lange, Rostock

Mathematische Logik gehört nicht zum Lehrplan der Ingenieure und auch nicht zu ihrem Aufgabenbereich. Aber logisches Denken gehört unbedingt zur Ingenieurkunst. Die Forderung, daß die Fachsprache einer Disziplin logisch sein soll, ist eigentlich selbstverständlich. Leider ist diese Forderung nicht immer erfüllt. Nachfolgend soll auf ein solches Beispiel hingewiesen werden, auf den Modellbegriff.

- Im Englischen bedeutet "model" ein Modell, ein Muster, ein Vorbild oder auch eine Vorführdame.
- Im Lateinischen bedeutet „modulus" Maß, Maßstab, Weise, Lied.

Nachfolgend soll betrachtet werden, was „Modell" in der Systemtechnik bedeutet.

■ Zunächst ist festzustellen, daß es in jeder Sprache mehrdeutige Ausdrücke gibt, Homonyme genannt;

z.B. bedeutet das Wort Schloß entweder eine Türverriegelung oder ein Gebäude (mit gleichem Artikel!) oder der Kunde und die Kunde (mit unterschiedlichen Artikeln). Was im Einzelfall gemeint ist, erkennt man leicht aus dem Kontext. Allerdings bereitet diese Mehrdeutigkeit bei einer automatischen Sprachübersetzung mit Hilfe eines Computers schon erhebliche Schwierigkeiten. Aber die künstliche Intelligenz befindet sich ja noch in starker Weiterentwicklung und wird auch dieses Problem lösen. In Form einer Verkopplung mit der menschlichen — der echten — Intelligenz als „Mensch-Computer-Tandem" bahnt sich bereits eine Lösung hierbei an.
Dagegen erscheint es als unzulässig, daß man eine Bezeichnung sowohl in einem Sinne oder im entgegengesetzten Sinne benutzt. Leider ist das beim Begriff „Modell" der Fall, wobei in dem einen Sinne ein Vorbild oder Original oder in dem entgegengesetzten Sinne ein Abbild (= Bild) gemeint ist, und zwar nicht in unterschiedlichen sondern im selben Fachgebiet, der Systemtheorie und der Systemtechnik, die als Systemwissenschaft zusammengefaßt werden.
Daneben stehen sprachlich eng verwandte Begriffe 'Model und Modul' die ihrerseits als Synonyme zueinander und als Homonyme in der Anwendung auf verschiedenen Fachgebieten benutzt werden.

■ Synonym bedeutet, daß zwei Begriffe sinnverwandt sind.

1. Anwendung des Modell-Begriffs in der Systemanalyse: Modell als Abbild

In der Systemwissenschaft ist man bestrebt, verschiedenartige Disziplinen unter einheitlichen Begriffen zusammenzufassen. Hierbei unterscheidet man
Systemanalyse und Systemsynthese.

- Die Physik als Grundlage der technischen Wissenschaften, die Medizin und andere Naturwissenschaften beschäftigen sich vor allem mit der Systemanalyse, um anorganische oder organische Vorgänge in ihrem Wirkungsablauf qualitativ und quantitativ zu untersuchen und die kausalen Zusammenhänge klarzustellen.
  Die Beschreibung von Zeitvorgängen (Ursache-Wirkungs-Relationen) wird stets an Hand eines vereinfachten Modells vorgenommen, das ein Abbild darstellt.
- In der theoretischen Physik und in der technischen Kybernetik wird ein dynamischer Prozeß durch ein System von Differentialgleichungen beschrieben, das man als
  — mathematisches Modell — bezeichnet.

Hierbei kann nicht jede Einzelheit des Prozesses beschrieben werden, sondern nur das Wesentliche des Prozesses. Es wird also ein Informationsverlust in Kauf genommen. In technischen Disziplinen genügt dazu auch eine Skizze oder ein Schaltschema.

▲ Das Abbild ist also unvollständig!

Aber ein derartiges Modell verschafft eine größere Klarheit und ermöglicht quantitative Aussagen über den Prozeßablauf.
Den Gegenstand zum Analysemodell bildet das Original, der betrachtete Gegenstand oder der ablaufende Prozeß (s. Bild 1). Dieser kann in keinem Fall als Modell bezeichnet werden, wenn eine Systemanalyse vorgenommen wird!

2. Anwendung des Modellbegriffes in der Systemsynthese: Modell als Vorbild

Ganz im Gegensatz zur Systemanalyse wird in der Systemsynthese der Modellbegriff benutzt:

■ Seit jeher wird in der Malerei und Bildhauerei, also der bildenden Kunst, das eigentliche Objekt, das abgebildet werden soll, als das Modell bezeichnet, z.B. eine menschliche Gestalt.

Je nach der herrschenden Kunstauffassung von Rubens bis Picasso kann das Bild sehr unterschiedlich ausfallen, besonders wenn es sich um eine Karikatur handelt. Auch hier wird oft ein Informationsverlust in Kauf genommen bzw. er wird sogar beabsichtigt, um wesentliche Merkmale hervorzuheben.
Aber keinesfalls wird das Abbild als Modell bezeichnet, sondern nur das Original.
Sofern es sich um völlig verschiedene Disziplinen mit eigenen Begriffsdefinitionen handelt, kann man gegen eine unterschiedliche Ausdrucksweise nichts einwenden.
Anders dagegen, wenn es sich um das gleiche Arbeitsgebiet und das gleiche Problem handelt! Dieser Fall liegt leider hier vor.

■ Die technischen Wissenschaften beschäftigen sich vorwiegend mit der Systemsynthese, genauer gesagt, mit der Konstruktion von Gegenständen aller Art. In diesem Fall stellt nach dem derzeitigen Sprachgebrauch das Modell das Muster oder den Bauplan zur Realisierung des Gegenstandes dar.

So spricht man im Automobilbau von einem neuen Automodell für eine zukünftige Serie von Fahrzeugen. Hierbei ist die Konstruktion bis auf alle Einzelheiten genau festgelegt.

Das „Modell" ist hier das Vorbild.

- Das Nachbaumuster, die technische Realisierung, soll hierbei das Original möglichst vollständig wiedergeben, jeder Informationsverlust bedeutet einen Mangel und muß vermieden werden (Bild 1).

Es zeigt sich somit, daß innerhalb des abgeschlossenen Fachgebiets, der Systemwissenschaft oder Systemtechnik, der Modellbegriff widersprüchlich benutzt wird.

▲ Ein Modell bedeutet in der Systemanalyse genau das Gegenteil als in der Systemsynthese.

In den Lexika wird durchweg nur das Modell der Systemanalyse definiert, vorzugsweise als mathematisches Modell.

3. Der Evolutionsprozeß von Synthesemodellen

Synthesemodelle sind stets einer Weiterentwicklung unterworfen, die man als Evolutionsprozeß bezeichnet. Weiterentwickelt wird stets in einer zweckbestimmten Richtung (z.B. Automobilbau zur Erhöhung der Geschwindigkeit, zur Verringerung des Kraftstoffverbrauchs oder des Schadstoffgehalts der Abgase, zur Verbesserung des Komforts u.a.).

- Hierbei wird jeweils ein neues Erstmuster (= Modell) angefertigt und dann als Serie nachgebaut.

Es existiert in solchem Fall ein Bauplan und ein realisiertes Muster. Es kann hiernach die Übereinstimmung der Nachbauexemplare in der Serienfertigung mit dem Muster (Modell) überprüft werden.

- In der Biologie begegnet man einem verwandtem Problem, der Evolution der Lebewesen, der Pflanzen und Tiere, jedoch mit einem wesentlichen Unterschied: es existiert kein Prototyp, kein Modell in Form eines Lebewesens.

Jedes Lebewesen hat eine begrenzte Lebensdauer und ist dann wieder als Einzelobjekt (als „Realisierung") verschwunden.
Vielmehr ist der Bauplan als Wachstumsanweisung (Wachstumsplan) in der Genstruktur als Erbanlage vorhanden und wird von einer Generation zur nachfolgenden Generation weitergegeben. Man hat inzwischen zwar der Natur das Geheimnis der Codierung der Erbanlagen entlockt, die aus der spezifischen Folge der 4 Nukleiden im Polynukleotid besteht. Man hat 1969 das Gen isoliert und 1970 synthesiert und man kann bereits als Gentechnik die Erbanlagen planmäßig verändern, also das Modell des Lebewesens. Aber das

große Geheimnis der Natur verbleibt, die *Decodierung* des chemisch übertragenen Wachstumplanes für das Lebewesen, das unendlich viel komplizierter aufgebaut ist als der Genübertragungsmechanismus mit der Doppelhelix. Das ist das noch ungelöste Rätsel der Natur, die modellgetreue Steuerung des Wachstums unter Realisierung aller Feinheiten der Art auf Grund der gespeicherten Information während der Ontogenese.

4. Artverwandte Begriffe

4.1. Model

Der Vollständigkeit halber sei erwähnt, daß die deutsche Sprache auch den Begriff „Model" kennt. Er leitet sich vom lateinischen Wort „modulus" ab und bedeutete ursprünglich ein kleines Maß. In der bildenden Kunst ist das eine altgriechische, von den Römern übernommene relative Maßeinheit zum Berechnen der Maßverhältnisse einer Säule: der Model beträgt die Hälfte des unteren Säulendurchmessers und wurde in 30 Partes (Teile) unterteilt.
Bis ins 18. Jh. hinein bezeichnete man damit auch eine Vorlage für Stickerei (Modelbücher). Ferner bedeutet Model in der Holztechnik einen durch Vorschnitt bearbeiteten Sägeblock zur Erzeugung von Kantholz, in der Lebensmitteltechnik ist es eine vertiefte Gußform zur Herstellung plastischer Gegenstände, z. B. für Spezialgebäck. Schließlich stellt ein Model in der Polygraphie eine in Holz gestochene Druckform für den Textilhanddruck dar.

4.2. Modul

Das ist eine Abkürzung von „modulus" (s. o.) und hat allein schon in der Mathematik unterschiedliche Bedeutung:
- als Betrag einer komplexen Zahl
- als Ziffernreihenfolge nach dem Komma bei Logarithmen
- als additiv geschriebene kommutative Gruppe (Abelsche Gruppe).

Im Maschinenbau bedeutet Modul die Kennziffer der Teilung bei Zahnrädern in mm, das ist das Verhältnis Teilkreisdurchmesser des Zahnrades zur Anzahl der Zähne (Teilung), geteilt durch π.
In der Elektronik bedeutet ein Modul ein bestimmtes Teilsystem einer Anlage, einen Baustein.
In der Computertechnik ist der Modul ein Programmbaustein.
Modul wird auch als Synonym für Model als Maß oder Maßstab verwendet, jedoch nicht für ein Modell.

Bild 1. Zur Doppeldeutigkeit des Modell-Begriffs

	Vorbild	Abbild
Modellbegriff in der Systemanalyse	Original (vollständig)	Modell z. B. vereinfachtes Schema oder Skizze bzw. mathematisches Gleichungssystem der Ursache-Wirkungs-Relation (unvollständig, mit Informationsverlust!)
Modellbegriff in der Systemsynthese	Modell z. B. Konstruktionsvorschrift oder Erstmuster (vollständig, ohne Informationsverlust); in der Biologie: Erbanlage, Wachstumsmodell der Gene	Nachbaumuster, technische Realisierung (auch möglichst vollständig, jedoch oft mit kleinen Mängeln); in der Biologie: Generationsfolge der Tiere und Pflanzen

NaA 113

Prof. em. Dr. sc. techn. Dr.-Ing. e.h. *Franz-H. Lange*, Ernst-Heydemann-Str. 3, Rostock, 2500

Impedanz- und Symmetrieverhalten der analogen Teilnehmeranschlußschaltung

K. F. Müller, Machern

Mitteilung aus dem VEB Nachrichtenelektronik „Albert Norden" Leipzig

Die Realisierung analoger Teilnehmeranschlußschaltungen für digitale Vermittlungssysteme tendiert zur Integration der BORSCHT-Funktionen in möglichst wenige integrierte Schaltkreise. Die Anforderungen an die Halbleitertechnologien sind erheblich, vor allem weil sie so unterschiedlich sind. Die Probleme bestehen darin, daß die Interface-Schaltung einerseits Spannungen bis zu 180 V, Spitzenleistungen bis zu 6 W bei Rufaussendung und kurzzeitige Spitzenströme von mehreren Ampere bei Überspannungseinwirkung verarbeiten muß; andererseits wird eine Gleichlaufgenauigkeit der Elementparameter im Promille-Bereich gefordert, wenn es um NF-Parameter wie Übertragungsdämpfung und Symmetrie geht. Daher findet man heute eine Trennung in Hochvolt- und Niedervoltschaltkreise. Den integrierten SLIC-Schaltkreis wird es daher möglicherweise nicht geben. Den grundsätzlichen Aufbau einer modernen analogen Teilnehmeranschlußschaltung zeigt Bild 1.

1. Aufbauprinzipien der integrierten SLIC-Schaltung

Die Anschlußschaltung nach Bild 1 wird für den Fall der Niederfrequenzübertragung betrachtet. Sie muß hinsichtlich ihrer NF-Übertragungseigenschaften 2 Aufgaben erfüllen: die Gabelfunktion, d. h. die Vierdraht-Zweidraht- und Zweidraht-Vierdraht-Übertragung einschließlich Richtungstrennung (Echokompensation), und die Erzeugung einer definierten Impedanz zur Zweidrahtleitung. Beide Aufgaben werden durch je eine Rückkopplungsschleife gelöst — Bild 2.

Die Schleife mit dem Übertragungsweg i erzeugt die Zweidrahtimpedanz. Die Schleife mit dem Übertragungsweg e dient der Kompensation des Empfangssignals, das über den Weg ER → Z → SR in die Senderichtung gelangt (Hohlklangeffekt, Schwingneigung). Ob der AD-DA-Wandler an der im Bild gezeichneten Stelle oder an der Stelle $1-1'$ und $2-2'$ angeordnet ist, bleibt für die Betrachtungen belanglos. In [1] wird die Impedanzschleife so in 2 Pfade aufgeteilt, daß der Speiseschaltkreis für sich allein eine Impedanz von 600 Ω reell hat. Komplexe Impedanzen werden durch einen zweiten, parallelliegenden Pfad im digitalen Signalprozessor (SLAC) ermöglicht. Diese Lösung ist besonders günstig, weil sie zu einem Speiseschaltkreis führt, der bei geringeren Anforderungen ohne Digitalprozessor einsetzbar ist.

Im Bild 3 werden 4 grundsätzliche Möglichkeiten der Impedanzsynthese und ihre Zweipolersatzschaltungen gezeigt. Unterschieden wird nach Speiseart und Gegenkopplung. Speiseart bedeutet, daß die Zweidrahtleitung entweder aus gesteuerten Spannungsquellen

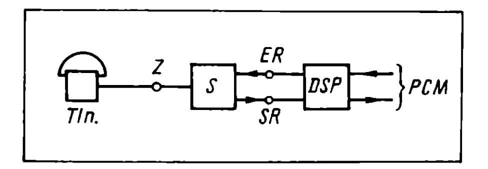


Bild 1. Grundsätzlicher Aufbau einer analogen Teilnehmeranschlußschaltung
Tln Teilnehmerapparat; Z Schnittstelle zur Zweidrahtleitung; S Speiseschaltung; ER, SR Empfangs- und Senderichtung; DSP zugeschnittener Signalprozessor

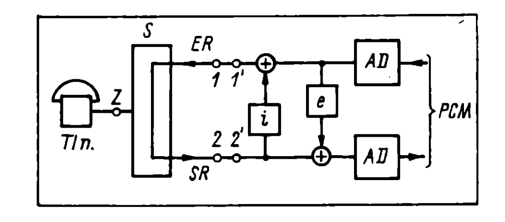


Bild 2. Die beiden Rückkopplungsschleifen der SLIC-Schaltung zur Impedanzerzeugung (i) und Echokompensation (e)

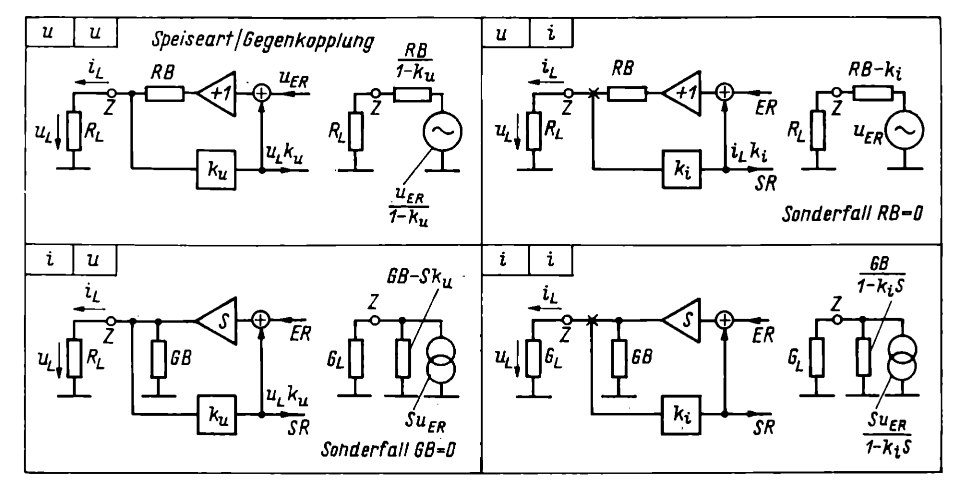
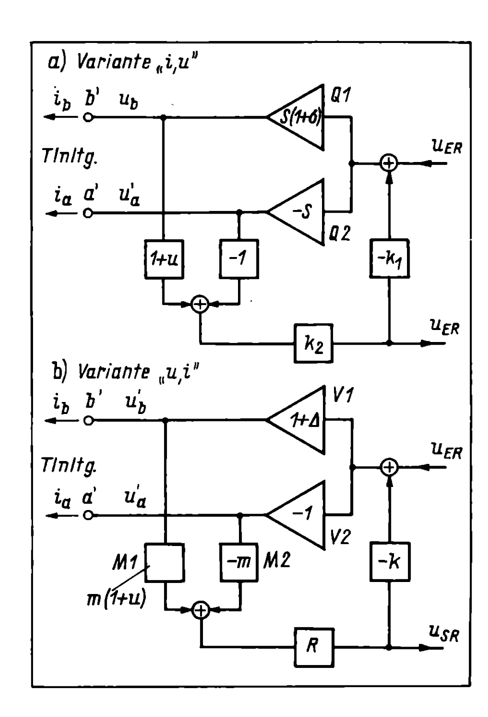
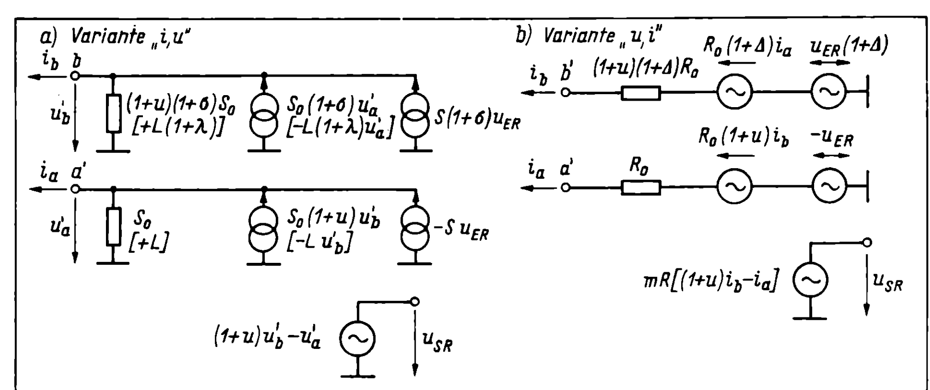


Bild 3. Möglichkeiten der Impedanzsynthese, unsymmetrische Darstellung, unterschieden nach Speiseart und Gegenkopplung; ku dimensionslos, ki in Ω

Bild 4. Ausführlichere Blockschaltbilder der SLIC-Varianten mit a) Stromspeisung/Spannungsgegenkopplung und b) Spannungsspeisung/Stromgegenkopplung
Die griechischen Buchstaben kennzeichnen relative Toleranzen der Baugruppen zueinander. ▶

Bild 5. Ersatzschaltungen der SLIC-Varianten nach Bild 4
Die Ausdrücke in rechteckigen Klammern gelten bei Einsatz einer Längsimpedanzschleife nach Bild 6.





Tafel 1. Transversal- und Longitudinalimpedanzen der SLIC-Varianten nach Bild 4
Die Werte in rechteckigen Klammern gelten, wenn die Längsimpedanzschleife nach Bild 6 eingeführt ist.

Variante	i, u		u, i
Transversalimpedanz			
$Z_{a'b'}$	$1/S_0$		$4\,R_0$
Longitudinalimpedanz			
$Z_{a'\perp}$	$-\,1/\mu S_0$	[1/2 L]	$-\,\mu R_0$
$Z_{b'\perp}$	$1/\mu S_0$	[1/2 L]	$+\,\mu R_0$

(u) oder gesteuerten Stromquellen (i) getrieben wird. Die Ausgangsspannung des Gegenkopplungspfades ist entweder der über der Zweidrahtleitung stehenden Spannung (u) oder dem in dieser Zweidrahtleitung fließenden Strom (i) proportional. In der Praxis durchgesetzt haben sich die Lösungen „u, i" [2] [3] [4] und „i, u" [1] [5], und zwar für die Sonderfälle RB = 0 bzw. GB = 0 nach Bild 3. Dies hat den Vorteil, daß die Impedanzen den Gegenkopplungsfaktoren direkt proportional sind und daß infolge RB = 0 bzw. GB = 0 bei Speisung der Teilnehmerleitung keine Leistungsverluste durch impedanzerzeugende Mittel auftreten.

2. SLIC Ersatzschaltungen der Varianten „i, u" und „u, i"

Ausführlichere Blockschaltbilder der beiden Varianten sind im Bild 4 dargestellt. Da hier vor allem Symmetriefragen zur Debatte stehen, sind die Übertragungsfaktoren der Baugruppen mit relativen Toleranzen zueinander angegeben. Im Bild 4a wurden beide Adern a, b der Teilnehmerleitung aus 2 spannungsgesteuerten Stromquellen Q 1 und Q 2 (Steilheit S) gespeist. Eine Spannungsmeßeinrichtung M1-M2 mißt die Differenz beider Adernspannungen (Transversalspannung). Das kann kein einfacher Differenzverstärker sein, weil die Adern a, b auf unterschiedlichen Gleichpotentialen liegen. Über den Pfad $-$ k1k2 (Operationsverstärker) wird der Impedanzkreis geschlossen.

Im Bild 4b wird die duale Anordnung mit Spannungsspeisung der Adern a und b durch die Verstärker V1 und V2 ($R_i = 0$) und Messung der Adernströme durch die Strommeßeinrichtung (Teile M1 und M2) dargestellt. Die Ausgangsströme der Teilmeßeinrichtungen sind ein Abbild (Maßstabsfaktor m, z.B. $m = 1/500$) der Adernströme. Ihre Summe ist bis auf einen Realisierungsfehler μ proportional zum Transversalstrom auf der Teilnehmerleitung. Über einen u-i-Wandler (Umsetzungswiderstand R) wird das Spannungssignal in Senderichtung gebildet und über den Gegenkopplungspfad $-\,k$ der dynamische Innenwiderstand des SLICs erzeugt.

Aus Bild 3 kann man für die Verknüpfung von Spannungen und Strömen folgende Gleichungen ableiten:

Variante „i, u" (Abkürzung $S_0 = \mathrm{k}S$)

$$i_b = -(1+\mu)(1+\sigma)\,S_0\,u_b' + (1+\sigma)\,S_0\,u_a' + (1+\sigma)\,S u_{ER} \quad (1a)$$

$$i_a = (1+\mu)\,S_0\,u_b' - S_0\,u_a' - S u_{ER} \quad (1b)$$

Variante u, i (Abkürzung $R_0 = \mathrm{mk}R$)

$$u_b' = -(1+\mu)(1+\Delta)\,R_0\,i_b + (1+\Delta)\,R_0\,i_a + (1+\Delta)\,u_{ER} \quad (2a)$$

$$u_a' = (1+\mu)\,R_0\,i_b - R_0\,i_a - u_{ER}. \quad (2b)$$

Aus diesen Gleichungen können beide Ersatzschaltungen nach Bild 5 abgelesen werden, die für die weiteren Symmetrieuntersuchungen wesentlich sind. Die Dualität beider Schaltungen wird hier sehr deutlich.

3. Transversal- und Longitudinalgrößen, Longitudinalschleife

Das Verhalten beider Schaltungsvarianten gegenüber Quer- und Längserregungen läßt sich nun anhand der Ersatzschaltungen nach Bild 5 ableiten. Die Ausgangsgrößen an den Klemmen a′, b′ sind Summen bzw. Differenzen aus Longitudinal- und Transversalgrößen

$$X_{a}' = X_t + X_l \quad (3a)$$

$$X_{b}' = -\,X_t + X_l. \quad (3b)$$

Längsgrößen werden entweder durch Unsymmetrien im System SLIC-Zweidrahtleitung oder durch Beeinflussung von außen induktiv erregt, z.B. durch zur Zweidrahtleitung parallel geführte Energieleitungen. Äußere Störer mit großer EMK werden daher meist hochohmig in die Zweidrahtleitung eingekoppelt.

Tafel 1 gibt einen Überblick über die Impedanz zwischen den Adern (Transversalimpedanz) und von Ader nach Masse (Longitudinalimpedanz). Für ideale SLICs sind die Impedanzen gegenüber Längsgrößen bei Stromspeisung unendlich und bei Spannungsspeisung Null. $-$ Daher schluckt die Variante u, i mit Spannungsspeisung die Längsströme über den Verstärkerinnenwiderstand problemlos,

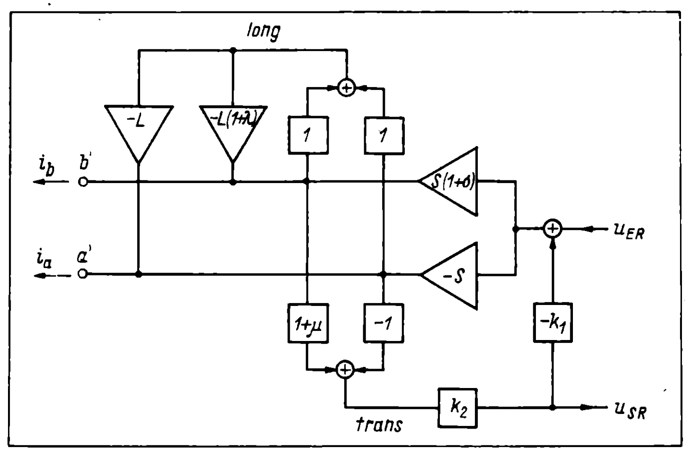


Bild 6. Die Schaltung nach Bild 4 ist um die Schaltung zur Erzeugung einer niedrigen Longitudinalimpedanz ergänzt worden

Tafel 2. Sende- und Empfangssymmetrie der SLIC-Varianten „i, u" und „u, i", erstere auch mit Längsimpedanzschleife
(Bedingungen: Anpassung und Pegelgleichheit)

Variante	i, u	u, i
Sendesignalsymmetrie	$20 \lg \frac{4}{\lvert\sigma\rvert}$	$20 \lg \frac{4}{\lvert\Delta\rvert + \lvert\varrho\rvert \frac{R_V}{R_V + Z/2}}$
mit Längsimpedanzschleife $[LR_V \to \infty]$	$20 \lg \frac{4}{\lvert\varrho\rvert \frac{R_V}{R_V + Z/2}}$	–
Empfangssignalsymmetrie	$20 \lg \frac{2}{\lvert\mu\rvert}$	$20 \lg \frac{2}{\lvert\mu\rvert + \lvert\varrho\rvert \frac{R_V}{R_V + Z/2}}$
mit Längsimpedanzschleife $[L \gg S_0]$	$20 \lg \frac{2}{\lvert\lambda\rvert + \lvert\varrho\rvert \frac{R_V}{R_V + Z/2}}$	–

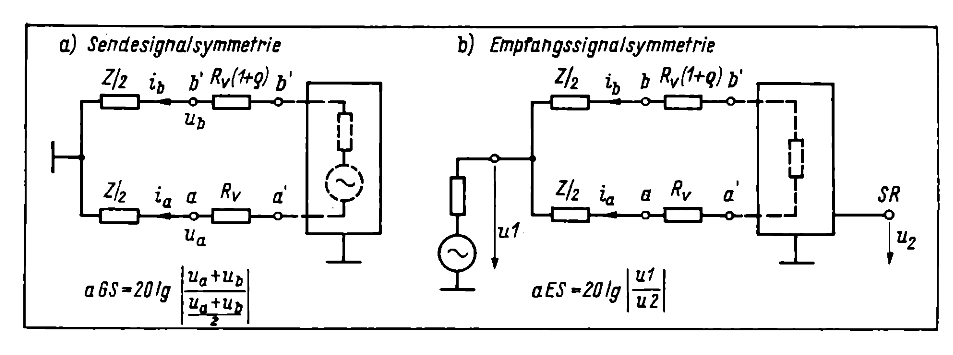


Bild 7. Symmetriemeßschaltungen nach CCITT 0.121
a) SLIC als Sender für Longitudinalsignale
b) SLIC als Empfänger für Longitudinalsignale

sofern er klein genug ist. Die Variante „i, u" dagegen benötigt noch eine zusätzliche Gegenkopplung, die für Längsgrößen entsprechend niedrige Innenwiderstände $Z_a'{}_\perp$ und $Z_b'{}_\perp$ erzeugt [4]. Die Stromquellenvariante wäre andernfalls durch sehr geringe Längsströme übersteuerbar. Sie würde durch Anstoß an die Gleichpotentialgrenzen ihre Linearität einbüßen. Bild 6 zeigt eine Möglichkeit der Längsimpedanzerzeugung [4]. Zwei zusätzliche Stromquellen werden von der Längsspannung ($U_a' + U_b'$) gesteuert. Die Steilheiten L der Quellen tolerieren gegeneinander um den Wert λ.

Die Analyse der Schaltung nach Bild 6 führt auf die Ergänzungen, die in der Ersatzschaltung nach Bild 5a in rechteckige Klammern gesetzt wurden. Die Längsimpedanzschleife wirkt sich auf die Transversalimpedanz $Z_a'{}_b'$ nicht aus. Für die Longitudinalimpedanzen $Z_a'{}_\perp = Z_b'{}_\perp = 1/2\,L$ können jedoch bei entsprechender Wahl der Steilheit L sehr kleine Werte realisiert werden.

4. Symmetrieforderungen nach CCITT 0.121

CCITT gibt in der Empfehlung 0.121 vier Symmetriemeßschaltungen an, die das Impedanz-Symmetrieverhältnis, die Signalsymmetrien von und zur Zweidraht-Teilnehmerleitung und das Gleichtaktunterdrückungsverhältnis betreffen. Die Schaltungen zur Messung der Signalsymmetrien vom und zum SLIC zeigt Bild 7a. Die Vorwiderstände R_V im Bild 7 sind unbedingt nötig für die Funktion der SLIC-Schaltung. Sie dienen u.a. der Überstrombegrenzung bei Blitzschlag. Da ihre Toleranz zueinander die Symmetrie beeinträchtigt, müssen sie berücksichtigt werden (Die SLIC-Schaltung hat demzufolge nur noch die Impedanz $Z - 2R_V$ zu synthetisieren).

An die Anschlüsse a', b' sind die Ersatzschaltungen nach den Bildern 5 und 6 angeschaltet zu denken. Um vergleichbare Ergebnisse zu erhalten, werden die Bedingungen „Anpassung" ($Z_i = Z_a = Z$) und Pegelgleichheit ($|U_{ab}| = |U_{SR}| = |U_{ER}|$) gesetzt. Daraus folgen $S_0 = S/2 = 1/(Z - 2R)$, $k_1 k_2 = 0.5$, $k_2 = Z/(Z - 2R_V)$ bzw. $4R_0 = 4kmR = Z - 2R_V$, $mR = Z/2$. Die damit errechneten Symmetrieverhältnisse sind in Tafel 2 zusammengefaßt. Es werden nur Fehler erster Ordnung berücksichtigt.

5. Sendesignalsymmetrie

In der reinen i, u-Variante werden die toleranzfreien Meßwiderstände von Strömen durchflossen, deren gegenseitige Toleranz von der Toleranz σ der Steilheit S abhängt. Wegen der Stromquellenspeisung hat die Toleranz ϱ der Vorwiderstände R_V keinen Einfluß auf die Spannung zwischen den Punkten a und b und damit auf die Generatorsignalsymmetrie.

Eine zusätzlich angebrachte Längsimpedanzschleife sorgt für exakt gegenphasige Aussteuerungen an den Punkten a', b', weil Unsymmetrien durch die Meßeinrichtung an dieser Stelle erfaßt werden (Spannungsnull). Daher beeinflußt die Toleranz ϱR_V der Vorwiderstände allein die Symmetrie der Spannungen über den Meßwiderständen. Diese absolute Toleranz ϱR_V im Zweig b'-Masse geht dabei auf den Gesamtwiderstand $R_V + Z/2$ bezogen ein.

Bei der Variante u, i werden in die Punkte a', b' unsymmetrische Spannungen eingespeist, die zum Verstärkungsfehler Δ proportional sind. Hinzu kommen von der Toleranz ϱ abhängige Fehler, die wiederum auf den Zweigwiderstand $R_V + Z/2$ zu beziehen sind. In der Praxis liegt der Faktor $R_V/(R_V + Z/2)$ in der Größenordnung von 1/5 bis 1/25, so daß die Toleranz ϱ der Vorwiderstände keinen bedeutenden Einfluß auf die Sendesignalsymmetrie hat. Die Toleranz μ der Meßeinrichtung geht nur mit 2. Ordnung in die Sendesignalsymmetrie ein.

6. Empfangssignalsymmetrie

Die Symmetrie der reinen Stromquellenvariante hängt nur von der Genauigkeit der Spannungsmeßeinrichtung ab. Wegen der hohen Längsimpedanz (vgl. Tafel 1) geht die Toleranz der Vorwiderstände nicht mit 1. Ordnung in das Symmetrieverhältnis ein. Eine zusätzliche Längsimpedanzschleife hält durch ihre Niederohmigkeit Längsgrößen von der Spannungsmeßeinrichtung fern, so daß deren Toleranz μ im Ausdruck für das Symmetrieverhältnis nicht erscheint. Die Ungenauigkeit der Stromquellen der Längsimpedanzschleife dagegen geht voll ein, während die Toleranz der Vorwiderstände auf die Zweigimpedanz $R_V + Z/2$ bezogen eingeht.

Der gleiche Aufbau der Symmetrieausdrücke der Variante „Stromquellenspeisung mit Längsimpedanzschleife" und bei der „Spannungsspeisung ist offensichtlich. Die Einführung der Längsimpedanzschleife verschiebt lediglich das Genauigkeitsproblem: die Toleranz μ der Meßeinrichtung wird durch die Toleranz λ der Stromquellen der Längsimpedanzschleife ersetzt.

NaA 34

Literatur

[1] *Rydin, A.; Sundvall, J.:* Line circuit component SLIC for AXE 10. Ericsson Review (1983) 4, S. 192–200

[2] *Sirsi, R. M.* u.a.: Integrated electronic subscriber line interface circuit with ringing on chip. IEEE J. of Solid-State-Circ. SC-18 (1983) 6, S. 665–677

[3] *Danneels, J. M.* u.a.: Monolithic 70 V bipolar linedriver IC for PCM SLIC. IEEE J. of Solid-State-Circ. SC-18 (1983) 3, S. 316–323

[4] *Borger, R.; Werner, H.-W.:* Integrierte Teilnehmerschaltung (SLIC) für digitale Vermittlungen. telcom report 8 (1985) 2. S. 102–107

[5] SLIC chip shrinks phone-line interfaces. Electronics (1985) 13, S. 54–57

Dr.-Ing. *Karl Friedrich Müller*, VEB Nachrichtenelektronik Albert Norden Leipzig, Melscher Str. 7, Leipzig, 7027

Dynamische Steuerung von Informationsflüssen in Kommunikationsnetzen unter dem Zuverlässigkeits- und Leistungsaspekt

D. Trstenský, Žilina; F. Beichelt, KDT; L. Sproß, KDT, Mittweida

Mitteilung aus der Hochschule für Verkehrs- und Nachrichtenwesen Žilina und der Ingenieurhochschule Mittweida

Die Leistungsfähigkeit von Kommunikationsnetzen kann bei gegebener Vermittlungs- und Übertragungstechnik im wesentlichen durch Vervollkommnung der topologischen Struktur des Netzes und durch Steuerung des Informationsflusses gesteigert werden. Zur Optimierung dieser Prozesse ist die Anwendung theoretisch fundierter Entscheidungsmodelle notwendig. Im folgenden werden einige neuere Grundmodelle vorgestellt.

Unter einem Kommunikations- bzw. Informationsnetz wird ein Netz zur Informationsübertragung und -verarbeitung verstanden. Mathematisch wird es durch einen ungerichteten zusammenhängenden Graphen $\underline{G} = (\underline{V}, \underline{E})$ mit der Knotenmenge $\underline{V} = 1, 2, \ldots, m$ und der Kantenmenge $\underline{E} = e_1, e_2, \ldots, e_n$ modelliert, der keine Schlingen und parallele Kanten enthält. In Fernmeldenetzen sind z.B. die Knoten Teilnehmer oder Vermittlungszentralen, während in Rechnernetzen die Knoten Rechner, Terminals bzw. Meßgeräte (bei Prozeßrechnern) sind. Jede Kante $e \in \underline{E}$ läßt sich in der Form $e = (i, j)$ mit $i, j \in \underline{V}$ darstellen. Im folgenden wird auch die Bezeichnung $e_{ij} = (i, j)$ verwendet. Jede Kante $e_{ij} \in \underline{E}$ entspricht stets einem Übertragungskanal zwischen den Knoten i und j. Existieren in der Praxis mehrere Kanäle zwischen i und j, so werden sie durch eine Kante e_{ij} verkörpert. Ohne Beschränkung der Allgemeinheit werden die Knoten als absolut zuverlässig vorausgesetzt, während die Kanten störanfällig sind. Die Indikatorvariablen X_i für den Zustand der Elemente e_i seien voneinander unabhängige Boolesche Zufallsgrößen:

$$X_i = \begin{cases} 1, & \text{wenn } e_i \text{ verfügbar ist,} \\ 0, & \text{wenn } e_i \text{ nicht verfügbar ist,} \end{cases}$$

$i = 1, 2, \ldots, n$. Die Zuverlässigkeit (Verfügbarkeit) $R(\underline{G})$ des Netzes ist dann nur eine Funktion der Verfügbarkeiten $p_i = P(X_i = 1)$ der Kanten e_i, $i = 1, 2, \ldots, n$.

1. Zuverlässigkeit von Kommunikationsnetzen

Für die zuverlässigkeitstheoretische Bewertung von Kommunikationsnetzen stehen je nach dem vordergründig interessierenden Anwendungsaspekt unterschiedliche Zuverlässigkeitskenngrößen zur Verfügung. Besondere praktische Bedeutung haben die paarweise Zusammenhangswahrscheinlichkeit $P(u, v)$ (zweier Knoten u und v) sowie die totale Zusammenhangswahrscheinlichkeit $P(\underline{V})$, die insbesondere für Fernmeldenetze die entscheidende Zuverlässigkeitskenngröße ist. $P(u, v)$ ist die Wahrscheinlichkeit dafür, daß eine Kommunikation zwischen den Knoten u und v möglich ist (d.h., es existiert zwischen u und v mindestens ein Weg, der nur aus funktionstüchtigen Kanten besteht). $P(\underline{V})$ ist die Wahrscheinlichkeit dafür, daß eine Kommunikation zwischen 2 beliebigen Knoten möglich ist (bez. der verwendeten graphen- und zuverlässigkeitstheoretischen Grundlagen s. [1]). Das Hauptproblem bei der Optimierung der topologischen Struktur von Kommunikationsnetzen bzw. ihrer Zuverlässigkeit besteht nicht in prinzipiell ungelösten theoretischen Problemen, sondern darin, daß der Rechenaufwand zur Ermittlung von Zuverlässigkeitskenngrößen mit wachsender Knoten- und Kantenzahl exponentiell wächst. So ist es gegenwärtig auch beim Einsatz der numerischen effektivsten Verfahren sowie modernster Rechner nicht möglich, Zuverlässigkeitskenngrößen für allgemein strukturierte Kommunikationsnetze mit mehr als 40 Knoten exakt zu berechnen (Ausnahmen bilden speziell strukturierte Netze, z.B. rekurrente sowie serien-parallel-reduzierbare Netze [4] [1]). Daher hat ein in [3] detailliert behandeltes Verfahren eine erhebliche praktische Bedeutung, das eine näherungsweise Berechnung der Zuverlässigkeit auch „großer" Kommunikationsnetze mit ausreichender Genauigkeit und in vertretbaren Zeiten ermöglicht. Es beruht auf folgenden Darstellungen der zu einem Kommunikationsnetz und der jeweiligen Zuverlässigkeitskenngröße zugehörigen Strukturfunktion

$$\varphi = \sum_{i=1}^{w} G_i$$

mit

$$G_i = \bar{A}_1 \bar{A}_2 \ldots \bar{A}_{i-1} A_i, \quad A_j = \prod_{\{i, i \in \underline{W}_j\}} X_i,$$

und

$$\varphi = 1 - \sum_{i=1}^{s} H_i$$

mit

$$H_i = \bar{C}_1 \bar{C}_2 \ldots \bar{C}_{i-1} C_i, \quad C_j = \prod_{\{i, i \in \underline{S}_j\}} \bar{X}_i.$$

Hierbei sind $\underline{W}_1, \underline{W}_2, \ldots, \underline{W}_w$ bzw. $\underline{S}_1, \underline{S}_2, \ldots, \underline{S}_s$ die minimalen Pfad- bzw. Schnittmengen des zugehörigen monotonen Systems. (Beispielsweise entsprechen im Fall der Berechnung von $P(u, v)$ die minimalen Pfadmengen den azyklischen Wegen von u nach v und im Falle der Berechnung von $P(\underline{V})$ des Gerüsten von $\underline{G}$.) Somit gilt

$$R(\underline{G}) = \sum_{i=1}^{w} E(G_i) = 1 - \sum_{i=1}^{s} E(H_i). \tag{1}$$

Die Berechnung von $E(G_i)$ und $E(H_i)$ erfolgt am besten durch Überführung der Booleschen Funktion G_i und H_i in Orthogonalformen [2] [8]. Für $a \leqq w$ und $b \leqq s$ ergeben sich hieraus die gewünschten Abschätzungen:

$$\sum_{i=1}^{a} E(G_i) \leqq R(\underline{G}) \leqq 1 - \sum_{i=1}^{b} E(H_i). \tag{2}$$

Im allgemeinen sind diese Schranken bereits für $a \ll w$ und $b \ll s$ hinreichend scharf. Dieser Effekt wird an folgendem Beispiel gezeigt.

Beispiel 1.

Bild 1 zeigt eine Netzstruktur mit $m = 3k + 1$ Knoten und $n = 5 + 7(k - 1)$ Kanten; $k = 1, 2, \ldots$. Die Kanten haben die gleiche Verfügbarkeit $p = 0{,}9$. Schon für $k = 6$ gibt es bezüglich $P(1{,}19)$ genau 55405 Minimalwege und 806 Minimalschnitte. Damit ist eine exakte Berechnung von $P(1, m)$ mit (1) und [8] für $k \geqq 6$ nicht mehr sinnvoll. Andererseits liefert aber (2) bereits für $a = 40$ Minimalwege bzw. $b = 12$ Minimalschnitte (jeweils kleinstmöglicher Mächtigkeit) gute Schranken für $P(1{,}19)$: $0{,}9765 \leqq P(1{,}19) \leqq 0{,}9976$. Bei der Schätzung von $P(\underline{V})$ mit (2) ergibt jedoch nur die obere Schranke eine brauchbare Näherung. Sie ist im Fall $k = 6$ schon für $b = 15$ völlig ausreichend ($= 0{,}9932$), da für wachsende $b > 15$ keine nennenswerte Bewegung in der oberen Schranke zu beobachten ist.

Es besteht also die Möglichkeit, die Zuverlässigkeit beliebig strukturierter Netze zumindest näherungsweise zu berechnen. Damit ist eine Grundvoraussetzung zur Schaffung solcher topologischer Struktur gegeben, die die vorgegebenen Zuverlässigkeitsforderungen erfüllen. Sind Kommunikationsnetze neu einzurichten bzw. zu vervollkommnen, so sind i.allg. Anzahl und Standort der Knoten vorgegeben. Freiheiten bestehen dagegen bei der Anzahl k von Übertragungskanälen zwischen 2 Knoten ($k = 0, 1, 2, \ldots$).

Davon hängen dann Zuverlässigkeit und Kapazität der Verbindung zwischen 2 Knoten ab. Bei entsprechenden Entscheidungen ist

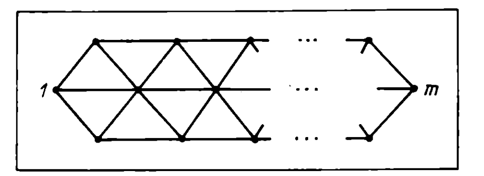
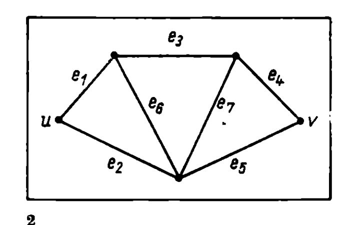
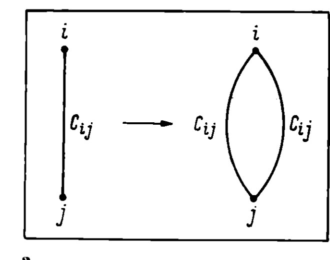


Bild 1. Netzstruktur variabler Größe



2

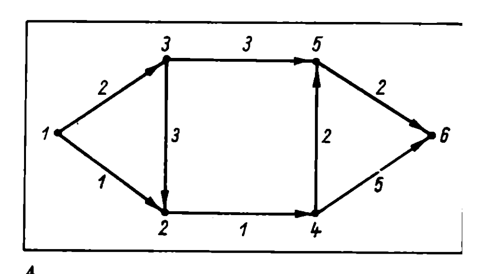


3

Bild 2
Netzstruktur mit 5 Knoten

Bild 3
Übergang von einer ungerichteten Kante zu einem Paar antiparalleler gerichteter Kanten

Bild 4
Gerichtete Netzstruktur mit bewerteten Kanten



4

die Kenntnis der zuverlässigkeitstheoretischen Wichtigkeit $I(i, \vec{p})$ einer vorhandenen bzw. zu installierenden Kante e_i von Bedeutung. In den meisten Fällen wird sie durch die Birnbaum-Importanz

$$I(i,\vec{p}) = \frac{\partial R(\underline{G})}{\partial p_i}, \; i = 1, 2, \ldots, n; \; \vec{p} = (p_1, p_2, \ldots, p_n),$$

gemessen [1]. Besonders interessant sind strukturelle Importanzmaße, da sie die zuverlässigkeitstheoretische Wichtigkeit eines Übertragungskanals nur aus der topologischen Struktur des Netzes ableiten, also nicht von möglicherweise unbekannten Kantenverfügbarkeiten abhängig sind. Diese Kriterien sind vor allem in der Phase der Projektierung von Kommunikationsnetzen bedeutungsvoll. Die strukturelle Birnbaum-Importanz von e_i ist definiert durch

$$I(i) = \frac{\partial R(\underline{G})}{\partial p_i}\bigg|_{p_1 = p_2 = \ldots = p_n = \frac{1}{2}}.$$

Beispiel 2.
Die in Bild 2 dargestellte Netzstruktur interessiert bezüglich $P(u,v)$. Es gilt

$$P(u, v) = p_3 R(\underline{G})_{e_3}) + \bar{p}_3 R(\underline{G}_{-e_3})$$

mit

$$\begin{aligned} R(\underline{G}_{e_3}) &= p_1 p_4 + p_2 p_5 - p_1 p_2 p_4 p_5 \\ &+ [p_1 p_5 + p_2 p_4 - p_1 p_2 p_4 - p_1 p_2 p_5 - p_1 p_4 p_5 \\ &- p_2 p_4 p_5 + 2 p_1 p_2 p_4 p_5] [p_6 + p_7 - p_6 p_7], \\ R(\underline{G}_{-e_3}) &= [p_2 + p_1 p_6 - p_1 p_2 p_6] [p_5 + p_4 p_7 - p_4 p_5 p_7]. \end{aligned}$$

Daher und aus Symmetriegründen betragen die strukturellen Birnbaum-Importanzen der Kanten $e_1, e_2, \ldots, e_7$:

$I(1) = I(4) = 0{,}2656$; $I(2) = I(5) = 0{,}4219$;

$I(3) = 0{,}1406$ und $I(6) = I(7) = 0{,}1005$.

Somit sind vor allem die Kanten e_2 und e_5 hinreichend zuverlässig auszulegen. Dieses Ergebnis ist nicht überraschend, da e_2 und e_5 den Weg zwischen u und v mit der kleinsten Kantenanzahl bilden.
Das bisher benutzte binäre Zustandsmodell erlaubt auch die Berücksichtigung unterschiedlicher Übertragungsleistungen (Kapazitäten) der Kanten. Dazu wird eine Kante genau dann als nicht verfügbar angesehen, wenn sie ausgefallen ist oder ihre Übertragungskapazitäten voll ausgeschöpft sind. Die Beschaffung der zugehörigen „Verfügbarkeiten" ist durch statistische Erhebungen, aber auch durch Simulation möglich.

2. Steuerung des Informationsflusses

Es seien $\Phi_{uv}(t)$ der zum Zeitpunkt t am Knoten u für den Knoten v bestimmte Verkehr, $\varphi_{uv}(t)$ die zugehörige Verkehrsintensität und $\Phi(t) = (\varphi_{uv}(t))$ die Matrix der Verkehrsintensitäten; $u, v = 1, 2,$ bis m. In Abhängigkeit vom Netzzustand, der von der Verkehrsbelastung sowie vom Störgeschehen im Netz abhängt, ist $\Phi(t)$ im Betrachtungszeitraum T zufälligen zeitlichen Schwankungen unterworfen:

$$\Phi(t_1) \to \Phi(t_2) \to \ldots \to \Phi(t_k) \to \ldots, t_1 < t_2 < \ldots < t_k,$$

$t_k \in T$. Änderungen des Netzzustandes können durch erfüllte bzw. neu entstandene Verkehrsaufgaben sowie durch Ausfälle von Netzelementen bzw. durch Abschluß ihrer Instandsetzung erfolgen. Im Fall der statischen Steuerung des Informationsflusses sind Knoten und Kanten so dimensioniert, daß mit hinreichend hoher Wahrscheinlichkeit die maximale Verkehrsintensität zwischen den Knoten u und v

$$\varphi_{uv}^{(\max)} = \max_{t \in T} \varphi_{uv}(t)$$

realisiert werden kann. Wesentlich effektiver, vor allem auch bei der Bemessung von Kommunikationsnetzen, ist die dynamische Steuerung des Informationsflusses [5]. Vom jeweils vorliegenden Netzzustand werden für neu entstandene Verkehrsforderungen die optimalen Routen der Informationsübertragung operativ ermittelt. Theoretisch kann nach jeder Änderung des Netzzustands die gesamt im Netz stattfindende Informationsübertragung neu optimiert werden. Solche Modelle sind allerdings rechentechnisch äußerst kompliziert (s. [7]) und technisch nur eingeschränkt realisierbar. Praktikabler ist, eine neu zu realisierende Informationsübertragung vom Knoten u zum Knoten v auf dem beim vorliegenden Netzzustand kürzesten verfügbaren Weg vorzunehmen. Es sei c_{ij} die Länge der Kante e_{ij}. Ist $\underline{W}_{uv}$ ein Weg von u nach v, so heißt

$$L(\underline{W}_{uv}) = \sum_{\{c_{ij}, e_{ij} \in \underline{W}_{uv}\}} e_{ij}$$

seine Länge. Speziell ist $\underline{W}_{uv}^{(\min)}$ ein kürzester Weg von u nach v, wenn

$$L(\underline{W}_{uv}^{(\min)}) = \min_{\underline{W}_{uv}} L(\underline{W}_{uv})$$

gilt.
Die Notwendigkeit, auf neu eintreffende Verkehrswünsche rasch zu reagieren, erfordert Algorithmen, die auch für größere Netze die jeweils kürzesten Wege möglichst schnell ermitteln. Von den bekannt gewordenen Verfahren zählt das von *Dijkstra* [6] mit zu den effektivsten. Es ist zunächst nur auf gerichtete Graphen anwendbar. Ein ungerichteter Graph kann jedoch bezüglich der Bestimmung von kürzesten Wegen in einen dazu äquivalenten gerichteten Graphen transformiert werden, wenn jede ungerichtete Kante e_{ij} durch 2 antiparallele gerichtete Kanten gleicher Länge c_{ij} $(= c_{ji})$ ersetzt wird (Bild 3). Ausgehend vom Knoten u ermittelt *Dijkstras* Verfahren die kürzesten Wege zu allen anderen Knoten $v \in \underline{V}$. In jedem Schritt des Verfahrens werden Marken $r_{\underline{M}}(x)$ für alle Knoten x auf folgende Weise festgelegt:

$r_{\underline{M}}(x)$ kürzeste Länge eines Weges von u zu x, wobei nur Zwischenknoten aus einer Menge $\underline{M}$ durchlaufen werden können.

Die Menge $\underline{M}$ variiert ebenfalls von Schritt zu Schritt. Es sei $x \notin \underline{M}$ ein Knoten mit der kleinsten Marke. Dann wird x in $\underline{M}$ aufgenommen und die Marken aller $y \notin \underline{M}$ werden neu durch

$$r_{\underline{M} \cup \{x\}}(y) = \min \{r_{\underline{M}}(y), r_{\underline{M}}(x) + c_{xy}\}$$

festgelegt. Gilt schließlich $\underline{M} = \underline{V}$, dann ist $r_{\underline{V}}(x)$ die Länge des gesuchten kürzesten Weges von u nach x. Zu Beginn werden $\underline{M} = \emptyset$ und $r_{\underline{M}}(x) = \infty$ gesetzt. Ferner wird $c_{xy} = \infty$ für $e_{xy} \notin E$ gesetzt. Im Algorithmus wird anstelle von $r_M(x)$ nur $r(x)$ geschrieben und eine Algol-ähnliche Notation verwendet [6].

Algorithmus von *Dijkstra*

```
Input:    Gerichteter Graph G = (V, E), c_ij > 0, u ∈ V.
Output:   Kürzeste Wege von u zu allen Knoten v ∈ V.
begin     set M: = u, r (u): = 0;
          for all y ∈ V \ {u} do r (y): = c_uy
          while M ≠ V do
          begin  find min {r (y); y ∉ M}; say r (x);
                 put M: = M ∪ {x};
                 for all y ∈ V \ M do
                 r (y): = min {r (y); r (x) + c_xy}
          end
end
```

Bild 4 zeigt einen gerichteten Graphen mit $m = 6$ Knoten. Die Kanten wurden mit ihren Längen bewertet. Die bei Anwendung des Algorithmus mit $u = 1$ auftretenden Mengen $\underline{M}$ sind $\emptyset$, $\{1\}$, $\{1, 2\}$, $\{1, 2, 3\}$, $\{1, 2, 3, 4\}$, $\{1, 2, 3, 4, 5\}$, und $\underline{V}$. Der kürzeste Weg zwischen $u = 1$ und $v = 6$ ist $e_{12}, e_{24}, e_{45}, e_{56}$. Seine Länge beträgt 6.

Fortsetzung auf Seite 131

Die quantisierte Fouriertransformation – eine Approximation der DFT

R. Sandau; E. Lorenz, Berlin

Mitteilung aus dem Institut für Kosmosforschung der AdW der DDR

Zur komprimierten Übertragung und/oder zur Interpretation von Signalen wird häufig eine Reduktion der Daten auf die problemspezifischen Merkmale angestrebt. Dabei muß neben der Datenreduktion auch die Merkmalrelevanz, hier genügt in den meisten Fällen die Translationsinvarianz, gewährleistet sein. Diese Eigenschaften der Merkmalextraktion setzen notwendigerweise eine nichtlineare Transformation bzw. eine lineare Transformation mit nichtlinearer Nachverarbeitung (z.B. Ermittlung des Leistungsspektrums) voraus [1] [2]. Es ist offensichtlich, daß nicht ein Merkmalraum bzw. eine nichtlineare Transformation für beliebige Signalklassen zuständig sein kann. Man wird sich bei der Suche nach relevanten Merkmalräumen auf bestimmte Signalklassen beschränken müssen. Für die Klasse der periodischen Signale bieten sich die Fourier-Leistungsspektren zur kompakten Beschreibung an. Hier erfolgt die Erzeugung der merkmalsrelevanten Daten durch Bildung der translationsinvarianten Spektralwerte aus den Ergebnissen der Fouriertransformation. In der digitalen Signalverarbeitung ist die Fouriertransformation jedoch eine sehr aufwendige Operation, die schon bei relativ geringen Echtzeitanforderungen den Einsatz von Mikroprozessoren bzw. von Mikrorechnern in Frage stellt.

In [3] wurden die Eigenschaften verschiedener verallgemeinerter Leistungsspektren bezüglich ihrer Relevanz und Kompaktheit im Vergleich zu dem DFT-Leistungsspektren untersucht. Als recht gute Approximation der DFT-Leistungsspektren können im Ergebnis dieser Untersuchungen die Leistungsspektren der quantisierten Fouriertransformation eingeschätzt werden.

In diesem Beitrag werden die Relevanz (zyklische Translationsinvarianz) und die Kompaktheit für die Leistungsspektren der DFT und der quantisierten Fouriertransformation mit 3 bzw. 5 Quantisierungsstufen (QFT3 bzw. QFT5) bei einer synthetisch erzeugten Signalform untersucht, die zeigt, wie sich eine sinusförmige Welle in einer CCD-Zeile an Bord eines Fernerkundungssatelliten unter Berücksichtigung der Sondierungsbedingungen (Sonnenstandswinkel, Atmosphäre, ...) abbilden würde (Bild 1a) [4].

In die Ausführungen wird ein Rechenzeitvergleich zur Durchführung der FFT und der FQFT3 mit bekannten Universalprozessoren bzw. -rechnern einbezogen.

1. Fouriertransformation DFT

Für eine reellwertige Folge

$$\{x(i)\} = \{x(0)\, x(1) \cdots x(N-1)\} \tag{1}$$

gilt die diskrete Fouriertransformation (DFT)

$$F(u) = \frac{1}{N} \sum_{i=0}^{N-1} x(i) . \exp[-2\,j\,\pi\, i\, u/N],\ i, u = 0, 1, \ldots, N-1 \tag{2}$$

mit

$$F(N/2 + l) = F^*(N/2 - l),\ l = 0, 1, \ldots, N/2 \tag{3}$$

(* symbolisiert den konjungiert komplexen Wert). Wegen (3) kann sich die Ausführung der DFT auf einen Halbraum beschränken. Außerdem wird man die DFT mit der Komplexität $0(N^2)$ nicht selbst berechnen, sondern einen schnellen Algorithmus (FFT) [5] der Komplexität $0(N \cdot \mathrm{ld}\, N)$ anwenden.

Weiterhin gilt für eine Folge

$$\{z(m)\} = \{x(m + h)\},\ h = 0, 1, \ldots, N-1 \tag{4}$$

der Verschiebungssatz

$$F_z(u) = W^{-uh} \cdot F_x(u),\ W = \exp[-2j\pi/N]. \tag{5}$$

Aus Gründen der Energieerhaltung gilt die Parsevalsche Beziehung für die DFT

$$\frac{1}{N} \sum_{i=1}^{N-1} [x(i)]^2 = \sum_{i=0}^{N-1} [F(u)]^2. \tag{6}$$

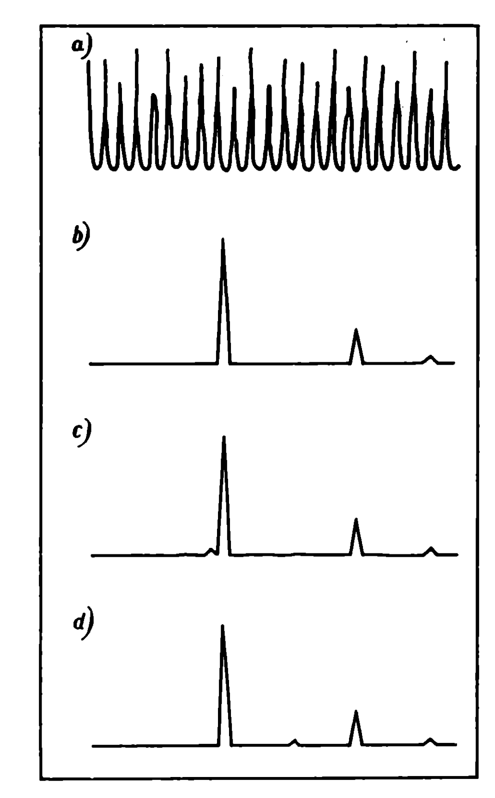


Bild 1. a) Abbildung einer sinusförmigen Welle im CCD-Sensor eines Fernerkundungssatelliten,
b) Fourier-Leistungsspektrum; c) QFT3-Leistungsspektrum;
d) QFT5-Leistungsspektrum

Fortsetzung von Seite 130

Die Parameter c_{ij} müssen nicht unbedingt als Längen der Übertragungskanäle e_{ij} interpretiert werden; sie können zum Beispiel auch Übertragungskosten bedeuten. Wird speziell $c_{ij} = -\ln p_{ij}$ gesetzt, wobei p_{ij} die Verfügbarkeit der Kante e_{ij} ist, so liefert der Algorithmus die zuverlässigsten Wege zwischen zwei Knoten. Im Fall $c_{ij} = 1$ ergeben sich die Wege mit der kleinsten Kantenanzahl.

NaA 33

Literatur

[1] *Beichelt, F.:* Zuverlässigkeit strukturierter Systeme. Berlin: VEB Verlag Technik 1988
[2] *Beichelt, F.; Sproß, L.:* An improved Abraham method for generating disjoint sums. IEEE Trans. Reliab. 36 (1987) 1, S. 70–74
[3] *Beichelt, F.; Sproß, L.:* Bounds on the reliability of binary coherent systems. IEEE Trans. Reliab. 38 (1989) 3, S. 346–348
[4] *Gadasin, V. A.; Ušakov, I. A.:* Nadežnost' složnych informacionno-upravljajuščich sistem. (Zuverlässigkeit infomationsgesteuerter Systeme) Moskva: Sovjetskoe Radio 1975
[5] *Lazarev, V. G.; Lazarev, Ju. V.:* Dinamičeskoje upravlenije potokami informacii v setjach svjazi. (Dynamische Steuerung von Informationsflüssen in Kommunikationsnetzen) Moskva: Radio i Svjaz 1983
[6] *Papadimitriou, Ch. H.; Steiglitz, K.:* Combinatorial Optimization: Algorithms and Complexity. New Jersey: Prentice-Hall 1982
[7] *Stassinopoulos, G. I.; Koukoutos, H.:* Optimale Steuerung beim Routewahlproblem unter besonderer Berücksichtigung spezieller Topologien. Wiss. Zeitschr. der Hochschule für Verkehrswesen „F. List", Sonderheft 33 (1988), S. 33–48
[8] *Sproß, L.:* Erzeugung von Orthogonalformen der Strukturfunktion binärer monotoner Systeme. Nachrichtentech., Elektron. 37 (1987) 9, S. 338–339

Doc. Ing. *Dušan Trstenský*, CSc., Hochschule für Verkehrs- und Nachrichtenwesen, Moyzesova 20, Čs-01088 Žilina; Prof. Dr. *Frank Beichelt;* Dr. sc. techn. *Lutz Sproß*, Ingenieurhochschule Mittweida, Platz der DSF 17 Mittweida, 9250

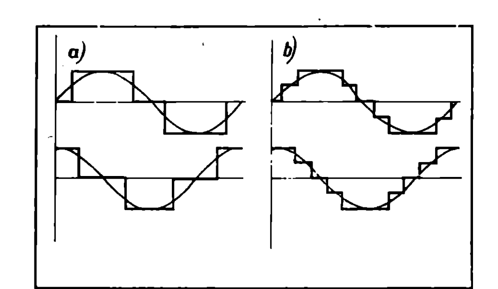


Bild 2. Quantisierung der Kreisfunktionen
a) mit 3 Niveaus
b) mit 5 Niveaus

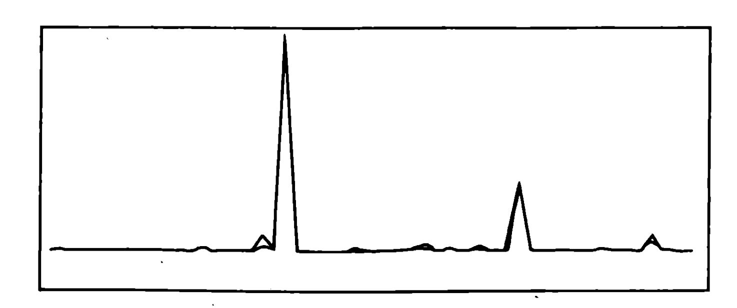


Bild 3. Maximum- und Minimum-Hüllkurven für die verschobenen QFT3-Leistungsspektren

Das Fourier-Leistungsspektrum $\underline{P}_F$ erhält man aus

$$\underline{P}_F = \underline{F} \cdot \underline{F}^*. \tag{7}$$

Wegen $W W^* = 1$ ergibt sich mit (5) das gleiche Leistungsspektrum für die verschobene Folge $\{z(m)\}$. Das Fourier-Leistungsspektrum ist somit invariant gegen zyklische Translation. Zum Vergleich mit den anderen, noch zu diskutierenden Transformationen zeigt Bild 1b das Fourier-Leistungsspektrum

$$\begin{aligned} P_F(0) &= |F(0)|^2 \\ P_F(n) &= 2\,|F(n)|^2,\ n = 1, 2, \ldots, N/2 - 1 \\ P_F(N/2) &= |F(N/2)|^2 \end{aligned} \tag{8}$$

der Abbildung einer sinusförmigen Welle im CCD-Sensor eines Fernerkundungssatelliten (s. Bild 1a) mit 8 bit Auflösung bei unterdrückten Mittelwert, d.h. $P_F(0) = 0$.

Zur Bestätigung der Translationsinvarianz und zur Erzeugung von Vergleichsmaßen für die Merkmalsrelevanz der anderen betrachteten Leistungsspektren wird das Fourier-Leistungsspektrum $\underline{P}_F$ des im Bild 1a dargestellten Signals mit dem Fourier-Leistungsspektrum $\underline{P}_F^{(l)}$ von l weiteren Realisierungen der Abbildungen sinusförmiger Wellen im CCD-Senser an Bord eines Fernerkundungssatelliten verglichen. Ein Maß für die Abweichungen der $\underline{P}_F^{(l)}$ von $\underline{P}_F$ ist der normierte mittlere quadratische Fehler

$$\mathrm{NMSE}(l) = \frac{\sum\limits_{i=0}^{N/2} [P_F(i) - P_F^{(l)}(i)]^2}{\sum\limits_{i=0}^{N/2} [P_F(i)]^2}, \quad l = 1, \ldots, \hat{l}. \tag{9}$$

Der Erwartungswert der Funktion NMSE(l) ergibt sich zu

$$E(\mathrm{NMSE}) = \frac{1}{\hat{l}} \sum_{l=1}^{\hat{l}} \mathrm{NMSE}(l) \tag{10}$$

und die Streuung zu

$$S(\mathrm{NMSE}) = \sqrt{\frac{1}{\hat{l}-1} \sum_{l=1}^{\hat{l}} \left([\mathrm{NMSE}(l)]^2 - \frac{1}{\hat{l}} \sum_{l=1}^{\hat{l}} [\mathrm{NMSE}(l)]^2 \right)}. \tag{11}$$

Mit $E(\mathrm{NMSE}) = 0{,}25 \cdot 10^{-14}$ und $S(\mathrm{NMSE}) = 0{,}62 \cdot 10^{-18}$ für $\hat{l} = 33$ kann die Translationsinvarianz als bestätigt betrachtet werden (die Werte liegen im Bereich der Rundungsfehler).

2. Quantisierte Fouriertransformation QFT

Auf der Suche nach Wegen, mit einfachen arithmetischen Mitteln ähnlich kompakte Merkmale wie aus der DFT zu erhalten, wird hier der Weg der Quantisierung gegangen, auf dem eine wachsende Anzahl von verarbeitungstechnisch günstigen Quantisierungsstufen für die DFT-Koeffizienten $W^{uk} = \cos(2\pi\, uk/N) - \mathrm{j} \cdot \sin(2\pi\, uk/N)$ zugelassen wird [6]. Im Bild 2a ist eine Möglichkeit gezeigt, bei der nur die 3 Quantisierungsniveaus -1, 0 und $+1$ auftreten. Eine weitere Möglichkeit ist im Bild 2b gezeigt. Dort werden 5 Quantisierungsniveaus ± 1, $\pm 1/2$ und 0 verwendet. Die Multiplikation mit den Werten $\pm 1/2$ kann dabei einfach durch Rechtsverschiebung des betreffenden Operanden erreicht werden. Bei der Einführung weiterer Quantisierungsniveaus sollte darauf geachtet werden, daß die Werte mit einfachen digitalen Strukturen (z.B. 2^{-i} mit Multiplexern) möglichst ohne Multiplizierer realisiert werden können.

Die Quantisierte Fouriertransformation kann als Approximation für die DFT aufgefaßt werden mit

$$Q(u) = \frac{1}{N} d_q(u) \sum_{k=0}^{N-1} x(k) \cdot \exp[\langle -\mathrm{j}\,\pi\, k\, u/N \rangle], \tag{12}$$

wobei $d_q(u)$ das Element eine diagonalen Normalisierungsmatrix $\underline{D}_q$ ist und der Ausdruck $\langle -\mathrm{j}\,2\pi\, ku/N \rangle$ die Quantisierung im o.g. Sinne symbolisiert. $\underline{D}_q$ ist nur zur Korrektur der Normierung der Transformation erforderlich, die im Ergebnis der Quantisierung gestört sein kann (im Unterschied zur Summe $\sin^2\varphi + \cos^2\varphi$ muß die Summe $\langle \sin\varphi \rangle^2 + \langle \cos\varphi \rangle^2$ nicht unbedingt $= 1$ sein).

2.1. Leistungsspektrum der QFT 3

Im Bild 1c ist das nach (8) ermittelte Leistungsspektrum für die QFT mit 3 Quantisierungsniveaus (P_{Q3}) im Vergleich zu anderen untersuchten Leistungsspektren dargestellt. Die Quantisierung wurde mit

$$\langle \sin\varphi \rangle = \begin{cases} 0, & 0 \leqq \varphi < 32{,}7^\circ \\ 1, & 32{,}7^\circ \leqq \varphi < 147{,}3^\circ \\ 0, & 147{,}3^\circ \leqq \varphi < 180^\circ \end{cases} \tag{13}$$

mit

$$\langle \sin\varphi \rangle = -\langle \sin(\varphi + \pi) \rangle \tag{14}$$

und

$$\langle \cos\varphi \rangle = \langle \sin(\pi/2 - \varphi) \rangle \tag{15}$$

so gewählt, daß die Fläche unter den Kreisfunktionen und den quantisierten Funktionen übereinstimmen. Für $\underline{D}_q$ wurde die Einheitsmatrix verwendet.

Bild 1c zeigt eine gute Übereinstimmung zum Fourierleistungsspektrum (Bild 1b). Die Kompaktheit der Beschreibung des Signals durch das QFT3-Leistungsspektrum ist mit der des DFT-Leistungsspektrums vergleichbar. Allerdings muß das QFT3-Leistungsspektrum noch auf wenigstens hinreichend gute zyklische Translationinvarianz überprüft werden.

Die im Bild 3 abgebildeten Maximum- und Minimum-Hüllkurven der QFT3-Leistungsspektren der um l Positionen verschobenen Signalfolge $\{x(m-l)\}$, $m = 0, 1, \ldots, N-1$, $l = 0, 1, \ldots, 33$ (33 ist die Korrelationslänge des Signals) zeigen eine gute Merkmalrelevanz.

Die qualitativ guten Ergebnisse werden auch bei der quantitativen Auswertung bestätigt. In Anlehnung an (9) gilt für die normierte quadratische Abweichung der QFT3-Leistungsspektren der Originalfolge $\{x(m)\}$ und der um l Positionen zyklisch verschobenen Folge $\{x(m+l)\}$

$$\mathrm{NMSE}_{Q3}(l) = \frac{\sum\limits_{i=0}^{N/2} [P_{Q3}(i) - P_{Q3}^{(l)}(i)]^2}{\sum\limits_{i=0}^{N/2} [P_{Q3}(i)]^2}, \quad l = 1, \ldots, \hat{l}. \tag{16}$$

Für $\hat{l}$ wird die Korrelationslänge der Folge $\{x(m)\}$ (d.h. $\hat{l} = 33$) genommen.

Für die Kreuzkorrelationsfunktion KKF_{Q3} zwischen den QFT3-Leistungsspektren der Folgen $\{x(m)\}$ und $\{x(m+l)\}$ gilt

$$\mathrm{KKF}_{Q3}(l) = \frac{\sum\limits_{i=0}^{N/2} P_{Q3}(i) \cdot P_{Q3}^{(l)}(i)}{\sqrt{\sum\limits_{i=0}^{N/2} [P_{Q3}(i)]^2 \cdot \sum\limits_{i=0}^{N/2} [P_{Q3}^{(l)}(i)]^2}}, \quad l = 1, \ldots, \hat{l}. \tag{17}$$

Tafel 1. NMSE und KKF der Leistungsspektren des verschobenen Signals

	Mittelwert FFT	QFT3	QFT5	Streuung FFT	QFT3	QFT5
NMSE	$.25 . 10^{-14}$	$.37 . 10^{-3}$	.0130	$.62 . 10^{-18}$	$.48 . 10^{-7}$	$.87 . 10^{-5}$
KKF		.9998	.9944	–	$.20 . 10^{-3}$	$.26 . 10^{-3}$
KKF_F		.9969	.9742	–	$.14 . 10^{-3}$	$.22 . 10^{-3}$

Entsprechend erhält man für die Kreuzkorrelationsfunktion KKF_{FQ3} zwischen den Leistungsspektren $P_F(i)$ der DFT und $P_{Q3}^{(l)}(i)$ der QFT3

$$KKF_{FQ3}(l) = \frac{\sum_{i=0}^{N/2} P_F(i) \cdot P_{Q3}^{(l)}(i)}{\sqrt{\sum_{i=0}^{N/2} [P_F(i)]^2 \sum_{i=0}^{N/2} [P_{Q3}^{(l)}(i)]^2}}, \quad l = 0, 1, \ldots, \hat{l}. \quad (18)$$

Die Erwartungswerte und die Streuungen der Kreuzkorrelationsfunktionen KKF_{Q3} und KKF_{FQ3} sowie der NMSE sind in Tafel 1 zusammengestellt.

2.2. Leistungsspektrum der QFT 5

Die 5 Quantisierungsniveaus wurden entsprechend den in [6] als optimal bezüglich der summarischen Abweichung der quantisierten Werte $\langle W \rangle$ von der Exponentialfunktion W angegebenen Verhältnissen

$$\langle \sin \varphi \rangle = \begin{cases} 0, & 0 \leqq \varphi < \pi/16 \\ 1/2, & \pi/16 \leqq \varphi < \pi/8 \\ 1, & \pi/8 \leqq \varphi \leqq \pi/2 \end{cases} \quad (19)$$

und deren Ergänzungen $\langle \sin \varphi \rangle = -\langle \sin(\varphi + \pi) \rangle$, $\langle \cos \varphi \rangle = \langle \sin(\pi/2 - \varphi) \rangle$ entsprechend (14) und (15) gewählt. Auch hier wurde, um Multiplikationen zu vermeiden, für die diagonale Normalisierungsmatrix $\underline{D}_q$ die Einheitsmatrix $\underline{I}$ eingesetzt.

Im Bild 1d ist das Leistungsspektrum P_{Q5} für das Signal aus Bild 1a im Vergleich zu den anderen Leistungsspektren dargestellt. Die Maximum- und die Minimum-Hüllkurven der um l Positionen verschobenen Signalfolge $\{x(m - l)\}$, $m = 0, 1, \ldots, N - 1$ und $l = 0, 1, \ldots, 33$ entsprechen etwa den im Bild 3 gezeigten.

Bei der quantitativen Auswertung ergeben sich die in Tafel 1 angegebenen Werte, wobei (16) bis (18) unter Verwendung des Indexes Q5 ausgewertet wurden.

3. Rechenzeitvergleich

Für die DFT bzw. für die QFT lassen sich schnelle, vom Grundtyp identische Algorithmen konstruieren [5] [6]. Die QFT eröffnet jedoch durch die sehr einfache Transformationsstruktur bei relativ geringem Informationsverlust gegenüber der DFT strukturelle Möglichkeiten (die Multiplikation mit den Koeffizienten 0, + 1, − 1 kann in arithmetisch-logischen Strukturen erfolgen), die sich bei entsprechendem Integrationsniveau von Spezialprozessoren entscheidend auf die Transformationsgeschwindigkeit und/oder die Volumen- und Leistungsverhältnisse niederschlagen können.

Aber auch bei Einsatz von Universalprozessoren bzw. -rechnern des „von Neumann"-Typs sind schon signifikante Geschwindigkeitsgewinne erzielbar. In Tafel 2 sind die entsprechenden Rechenzeiten für eine Signalfolge mit $N = 128$ reellen Stützstellen bei 8 bit- und 16 bit-Verarbeitung für die Mikroprozessoren U880 und U881/882 sowie U8000 zusammengestellt. Es ist zu erkennen, daß die Rechenzeitgewinne im Bereich 2 bis 3 liegen. Diese Zahlen zeigen aber auch, daß die Universalprozessoren mit der „von Neumann"-Architektur für komplexe Signalverarbeitungsaufgaben wenig geeignet sind, weil der Organisationsaufwand, der sequentiell zu den arithmetischen Operationen erfolgt, und die beschränkte Busbandbreite sich nicht umgehen lassen. Günstiger wäre hier natürlich der Einsatz digitaler Signalprozessoren der Harvard-Architektur oder gar der Einsatz von Spezialprozessoren mit einem entsprechenden Parallelisierungsgrad und/oder Grad der Pipelineverarbeitung in den Befehls- und Signalverarbeitungsstrukturen. Die in Tafel 2 angeführten Zeiten beziehen sich nur auf die reine Transformation; die Zeiten für die Bitspiegelung der Operandenadressen und die Ermittlung der translationsinvarianten spektralen Leistungen sind in diesen Zeiten nicht enthalten.

4. Zusammenfassung der Ergebnisse

Aus Bild 1 ist zu ersehen, daß die Leistungsspektren der quantisierten Fouriertransformation gute Approximationen des DFT-Leistungsspektrums liefern. Um die Ergebnisse quantifiziert vergleichen zu können, sind in Zeile 1 von Tafel 1 die Erwartungswerte und die Streuungen der normierten quadratischen Abweichungen NMSE der Leistungsspektren des um l Positionen verschobenen Signals ($l = 1, \ldots, 33$; 33 ist die Korrelationslänge des Signals) vom Leistungsspektrum des Originalsignals als Funktion des Transformationstyps (FFT, QFT3, QFT5) angegeben. Außerdem sind in Zeile 2

Tafel 2. Rechenzeiten für FFT und QFT3 mit Universalprozessoren für N-128 reelle Stützstellen

	8 bit-Verarbeitung		16 bit-Verabeitung	
	Z8	U880	U880	U8000
t_{FFT} in ms	84	106	378	33
t_{FQFT} in ms	40	52	115	18

der Tafel 1 die Erwartungswerte und die Streuungen der Kreuzkorrelationsfunktionen KKF zwischen den vom Transformationstyp abhängigen Leistungsspektren des Originalsignals und den entsprechenden Leistungsspektren des um l ($l = 0, \ldots, 33$) Positionen verschobenen Signals und in Zeile 3 die Erwartungswerte und die Streuungen der Kreuzkorrelationsfunktionen KKF_F zwischen dem Fourier-Leistungsspektrum und den Leistungsspektren der QFT3 und QFT5 des um $l = 0, \ldots, 33$ Positionen verschobenen Signals angeführt.

Die geringen Unterschiede in den Ergebnissen aus der quantisierten Fouriertransformation mit 3 bzw. mit 5 Quantisierungsniveaus (QFT3 bzw. QFT5) und die teilweise besseren Ergebnisse bei der QFT3 zeigen, daß diese Ergebnisse durch Variation der Schwellwerte und der Elemente der diagonalen Normalisierungsmatrix $\underline{D}_q$ sicherlich noch optimiert werden können. Legt man die vorhandenen Ergebnisse aus der Simulation zugrunde, kann die quantisierte Fouriertransformation mit 3 Quantisierungsniveaus QFT3 als gute Approximation der DFT betrachtet werden. Es müssen jedoch, verglichen mit den Ergebnissen aus der DFT, Abstriche bezüglich der Merkmalsrelevanz (zyklische Translationsinvarianz), der Kompaktheit der Merkmale (potentielle Kompressionsrate) und, wenn für die diagonale Normalisierungsmatrix $\underline{D}_q$ die Einheitsmatrix $\underline{I}$ eingesetzt wird, auch der Energieäquivalenz zwischen dem Signal und dem Transformationsergebnis gemacht werden.

Die in Tafel 2 zusammengestellten Rechenzeitanforderungen für die QFT3, die um den Faktor 2 bis 3 geringer sind als die Rechenzeiten für die Durchführung der FFT, erlauben es, den Einsatzbereich der Universalprozessoren vom von-Neumann-Typ für die Echtzeit-Merkmalextraktion zu erweitern. Es sei jedoch darauf hingewiesen, daß die Ergebnisse von Tafel 1 und Bild 1 von einer konkreten Signalform, einer Stichprobe aus der Klasse der periodischen Signale abgeleitet wurden. Sie müssen im Bedarfsfall anhand des konkreten Datenmaterials bzw. der aufgabenspezifischen Signalformen überprüft bzw. bestätigt werden.

NaA 13

Literatur

[1] *Müller, X.:* Schnelle, translationsinvariante Transformationen zur Bearbeitung digitaler Grauwertbilder. Dissertationsschrift, Universität Karlsruhe 1982. Fortschrittsbericht (Reihe 10, Nr. 17) der VDI-ZEITSCHRIFTEN, VDI-VERLAG, Düsseldorf 1982

[2] *Burkhardt, H.:* Transformationen zur lageinvarianten Merkmalsgewinnung. Habilitationsschrift, Universität Karlsruhe 1979. Fortschrittsbericht (Reihe 10, Nr. 7) der VDI-ZEITSCHRIFTEN, VDI-VERLAG, Düsseldorf 1979

[3] *Sandau, R.:* Transformationen und Prozessorstrukturen zur 2D-Merkmalsextraktion in der Fernerkundung. Diss., AdW der DDR, Forschungsbereich Geo- und Kosmoswissenschaften, Berlin 1987

[4] *Jahn, H.; Oertel, D.; Fellberg, G.:* Untersuchungen struktureller Parameter der Ozeanoberflächen mit optoelektronischen Sensoren auf Satelliten. Institut Kosmicheskich Issledovanij (IKI) Akademii Nauk SSSR 1986

[5] *Brigham, E. O.:* The Fast Fourier Transform. Engleweed Cliffs, New Jersey: Prentice Hall inc., 1974

[6] *Jareslavskij, L. P.; Mersljakov, N. S.:* Digitale Holographie. Moskau: Nauka 1982

Dr.-Ing. *Rainer Sandau*; Dipl.-Phys. *Eckehard Lorenz*, Institut für Kosmosforschung der AdW der DDR, Rudower Chaussee 5, Berlin, 1199

Arhythmische Signalcodierungen und Kompandierung ihrer Zeitbasis

W. Zühlke, Ilmenau

Mitteilung aus der Technischen Hochschule Ilmenau, Sektion Informationstechnik und theoretische Elektrotechnik

Während bei der bekannteren rhythmischen Abtastung bandbegrenzter Signale mit fester Abtastfrequenz und dadurch mit konstanten Abständen zwischen aufeinanderfolgenden Signalabtastungen gearbeitet wird (PCM, DM)[1]), treten bei signaladaptiv arbeitenden Abtastverfahren unregelmäßige Abstände zwischen den Abtastungen auf. Die Abtastungen und damit die zu übertragenden Informationen sind arhythmisch. Die Abtastzeitpunkte werden je nach Verfahren nach bestimmten Merkmalen (Phasen) im Signal, wie Nulldurchgänge des Signals oder von Ableitungen des Signals, oder nach bestimmten Aktivitäten im Signal oder in seinen Ableitungen, wie Verlassen vorgegebener Intervalle oder Gebiete, gewählt. Sie sind dadurch im Mittel seltener als bei rhythmischer Abtastung, bei der der Abtastabstand nach der im Signal (hin und wieder) vorkommenden höchsten Signalfrequenz festgelegt werden muß. Infolge der durch das Verfahren festgelegten und damit Sender und Empfänger bekannten Vereinbarungen enthält die zu übertragende Information bei signaladaptiver arhythmischer Abtastung viel weniger Redundanz als jene bei rhythmischer Abtastung, was soweit geht, daß bei einigen Verfahren nur noch das Vorzeichen der Änderung übertragen werden muß. Während bei rhythmischer Abtastung die Redundanz erst in den nachfolgenden Stufen (Transformation, Quantisierung und Codierung) reduziert werden kann, beginnt die Redundanzreduktion bei arhythmischen Verfahren schon bei der Signalabtastung.

Arhythmische Abtastwerte haben noch eine weitere bemerkenswerte Eigenschaft. Die zeitlichen Abstände der Abtastwerte und der daraus abgeleiteten Impulse oder Impulsgruppen sind informationstragend, insbesondere lassen die Abstände auf die momentane obere Grenzfrequenz schließen. Das läßt sich zur Kompandierung der Zeitbasis ausnutzen und damit zur Bandbreitenkompression für die Übertragung und Speicherung von Signalen, wie sie im folgenden beschrieben ist.

Zur Übertragung der codierten arhythmischen Information ist zu bemerken, daß die mit diesen Verfahren angestrebte und erreichte niedrigere mittlere Impulsrate bei frequenzmultiplexer Übertragung vorteilhaft ist und für Codemultiplexsysteme impulsarme Codierungen geradezu gefragt sind. Dagegen scheidet die zeitmultiplexe Übertragung im PCM-Rahmen aus, aber die Übertragung als komprimiertes adressiertes Paket in einem mit der Datenübertragung gemeinsamen paketvermittelnden Netz bietet sich an.

[1]) DM Deltamodulation

1. Arhythmische Verfahren

Auf die Vorteile der Extremwertabtastung (EWM) nach Bild 1 wurde schon in früheren Veröffentlichungen [3] [4] [5] hingewiesen. Daß sie dennoch wenig genutzt wurde, liegt daran, daß mit den bislang realisierbaren Interpolatoren (Halteglied, linearer Interpolator) keine ausreichende Qualität im Demodulator erreicht wurde und kompliziertere Interpolatoren zu aufwendig waren. Inzwischen hat nun die Technologie gewaltige Fortschritte gemacht, so daß auch der Cosinusinterpolator nach [5] Tafel 2 realisierbar wurde, weiterentwickelt werden konnte, und an kompliziertere Interpolationsalgorithmen, wie Spline- und Lagrangealgorithmen, gedacht werden kann. Außerdem gibt es Vorschläge, durch zusätzliche Stützstellen die Interpolation zu vereinfachen.

Für die weniger bekannte Amplitudendifferenzcodierung (DAC) nach Bild 2, (die als asynchrone Deltamodulation bezeichnet wird, aber weder asynchron noch eine Variante der Deltamodulation ist, wie unten erklärt wird) gibt *Steele* [2] einen Gewinn im Signal-Rausch-Verhältnis von 6 dB gegenüber der Deltamodulation (DM) an. Das hat nicht ausgereicht, die kompliziertere Realisierung in Kauf zu nehmen. Aber, erstens ist die Realisierung heute wesentlich billiger und zweitens kann auf Grund der entwickelten Interpolationstechnik die Amplitudenstufung jetzt viel günstiger gestaltet werden, so daß die Bitrate der DAC weiter sinkt und die Qualität des reproduzierten Signals steigt. Zur Erklärung des Verfahrens sind im Bild 2 in Amplitudenrichtung linear gestufte Schwellwerte eingezeichnet. Jedesmal, wenn das Signal einen benachbarten Schwellwert überschreitet, wird ein Impuls ausgegeben. Dieser Impuls muß zwei verschiedene Pegel haben, um dem Empfänger mitzuteilen, ob der höhere oder niedrigere benachbarte Schwellwert überschritten wurde bzw., ob die Signaländerung positiv oder negativ war. Es muß also lediglich das Vorzeichen der Signaländerung codiert werden und zwar im richtigen Zeitpunkt. Im Bild 3 ist eine günstigere logarithmische Stufung der Schwellwerte gezeigt. Die lineare DAC ist das amplitudenmäßige Analogon zu einer reinen Abtastung im zeitlichen Abstand DT und diese beiden Verfahren sind die Grenzverfahren der unten erklärten Distanzcodierung (DIC). Sie ergeben sich durch $a = 1; b = 0$ bzw. $a = 0; b = 1$ in (1a).

Die Distanzcodierung (DIC) [1] nach Bild 4 ist ein neues signaladaptives Abtastverfahren. Bei ihr werden die Abtastpunkte nicht äquidistant in der Zeitrichtung gewählt, wie bei rhythmischer Abtastung, und nicht äquidistant in der Amplitudenrichtung, wie bei der DAC-Grundvariante, sondern die Abtastpunkte werden in der DIC-Grundvariante äquidistant auf der Signalkurve ausgewählt. Die Wahl der Abtastpunkte paßt sich damit sowohl starken Amplitudenänderungen als auch schnellen zeitlichen Änderungen, also der Dynamik und der Frequenz an, und wählt dadurch in geeigneter Form die Stützpunkte (Gitterpunkte) für die Interpolation aus. Im Bild 4 sind für eine nichteuklidische Distanz DI nach

$$DI = a \cdot |DA| + b \cdot DT \qquad (1a)$$

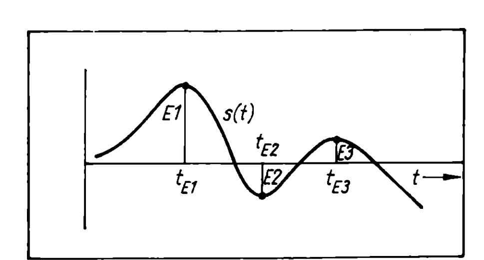


Bild 1. EWM-Abtastung

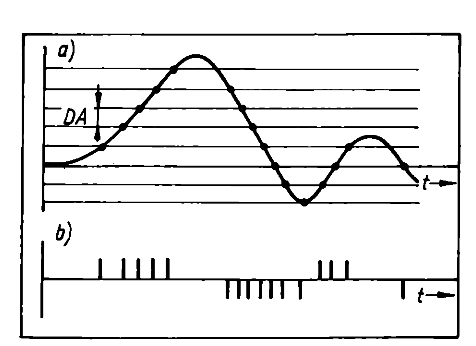


Bild 2. Amplitudendifferenzcodierung (DAC)
a) äquidistante Amplitudenschwellen
b) codierte Impulsfolge der DAC

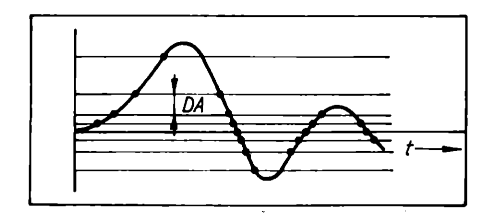


Bild 3. Logarithmisch gestufte Amplitudenschwellen der nichtlinearen DAC

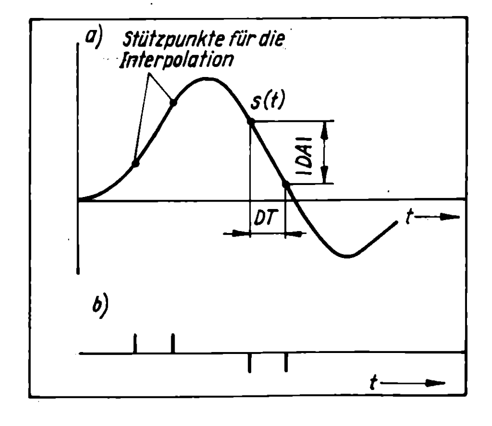


Bild 4. Distanzcodierung (DIC)
a) Stützpunktauswahl
b) codierte Impulsfolge der DIC

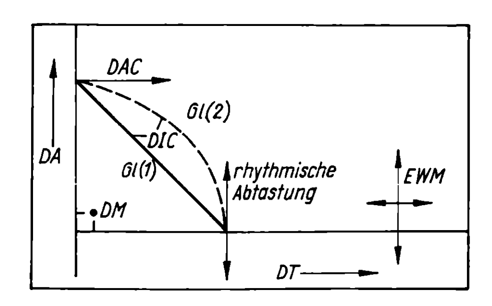
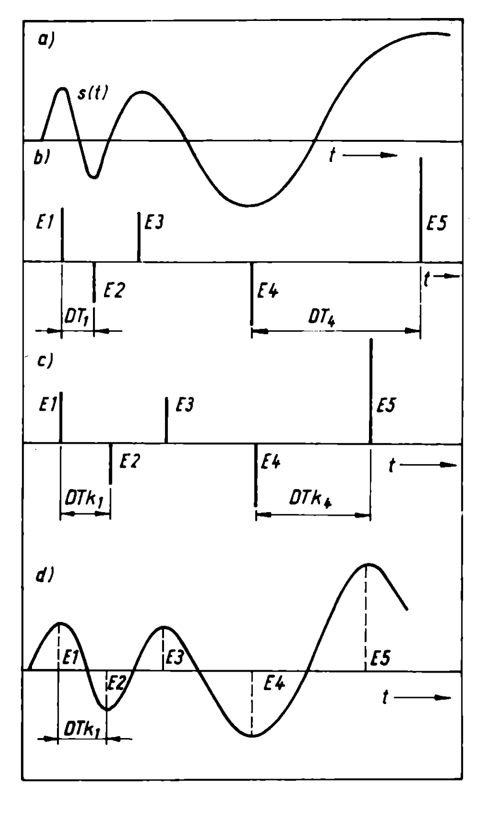


Bild 5. Einordnung der Verfahren



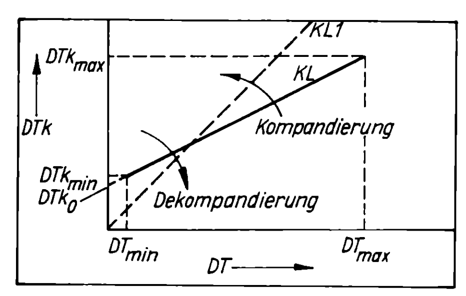


Bild 7. Lineare Kompanderkennlinie

Bild 8. Mögliche Kompanderkennlinien

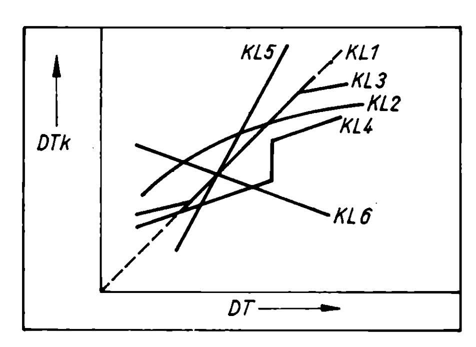


Bild 6. Kompandierung der Zeitbasis
a) Signalverlauf
b) Extremwertabtastfolge
c) zeitlich kompandierte Extremwertabtastfolge
d) analoge interpolierte zeitlich kompandierte Extremwertabtastfolge

die sich ergebenden Stützpunkte auf der Signalkurve gezeichnet. Jedesmal, wenn die sich aus (1a) ergebende Distanz DI einen vereinbarten Wert D erreicht, wird ein neuer Stützpunkt für die im Empfänger notwendige Interpolation ausgewählt. Für den Empfänger genügt es, im zeitlichen Abstand DT zum vorhergehenden Stützpunkt einen Impuls zu erhalten, der das Vorzeichen der Amplitudenänderung DA angibt. Der Betrag der Amplitudenänderung |DA| ist zwar variabel, kann aber nach (1b) berechnet werden, da D, a und b vereinbart sind und DT vom Empfänger gemessen werden kann.

$$|\mathrm{DA}| = (D - b \cdot \mathrm{DT})/a\,. \tag{1b}$$

Aus dem gemessenen DT, dem berechneten |DA|-Wert, dem übertragenen Vorzeichen von DA und den gespeicherten Koordinaten des vorhergehenden Stützpunktes läßt sich der neue Stützpunkt bestimmen. Auch mit euklidischer Distanz nach

$$D^2 = (a \cdot \mathrm{DA})^2 + (b \cdot \mathrm{DT})^2 \tag{2a}$$

und

$$|\mathrm{DA}| = \frac{1}{a} \sqrt{D^2 - (b \cdot \mathrm{DT})^2} \tag{2b}$$

läßt sich arbeiten. Man erhält damit am Sender eine Approximation des Signals durch einen „Polygonzug" mit Geradenstücken, die alle die gleiche Länge haben, mit Wahl der Stützpunkte in äquistanten Schritten auf der Signalkurve (nach [1] Bild 3). Die technische Realisierung dieser Gleichung ist softwaremäßig auch durch gespeicherte Tabellen möglich und damit sehr schnell. Weitere Ausgestaltungen des Verfahrens sind in [1] angegeben.

Wie bereits erwähnt, ergibt sich aus (1a) für $a = 0$ eine rhythmische Abtastvariante. Immer wenn $\mathrm{DI} = b \cdot \mathrm{DT}$ (mit b vorzugsweise $= 1$, genauer: $1 \cdot s^{-1}$) gleich dem Wert D ist, wird ein Stützpunkt gewählt. Da nun aber, wegen $a = 0$, die Amplitudendifferenz keinen Einfluß auf die Distanz hat, kann sie auch nicht am Empfänger berechnet, sondern muß ihm übermittelt werden.

Im Fall $b = 0$ wird immer, wenn $\mathrm{DI} = a \cdot |\mathrm{DA}|$ (mit a vorzugsweise $= 1 \cdot V^{-1}$) gleich dem Wert D ist, ein Stützpunkt gewählt, was dem DAC-Verfahren nach Bild 2 entspricht (bei linearer Amplitudenstufung). Im Bild 5 ist der Zusammenhang für die linearen Grundvarianten dargestellt.

2. Kompandierung der Zeitbasis

Bei der Einschätzung der arhythmischen Verfahren wurde immer vorsichtig formuliert, daß die Abtastungen im Mittel seltener sind als bei rhythmischer Abtastung. Tritt nun aber die obere Grenzfrequenz im Signal auf, müssen auch arhythmische Abtastverfahren sehr häufig abtasten und kommen zumindest kurzzeitig auf ähnliche Abtastgeschwindigkeiten wie die PCM und darüber (DIC). Dieser kritische Fall tritt bei der Extremwertabtastung im Bild 6b im linken hochfrequenten Teil des gezeigten Signalverlaufs auf. Innerhalb der kurzen Zeitspanne DT_1 soll die Information über die Amplitude des Extremwertes E_1 als Codewort übertragen werden. Im rechten, niederfrequenten Teil des Signals ist dagegen viel Zeit zur Übertragung des jeweiligen Codewortes vorhanden.

Für die Distanzcodierung und die Amplitudendifferenzcodierung gilt das Problem sinngemäß. Bei diesen ist zwar stets nur ein Impuls zu übertragen, aber dessen Dauer wird durch die kürzeste auftretende Zeitdifferenz $\mathrm{DT}_{\min}$ bestimmt und diese Dauer des Impulses bestimmt die notwendige Übertragungsbandbreite.

Um dieses Problem zu entschärfen und $\mathrm{DT}_{\min}$ zu vergrößern, wurde eine Dehnung (Expansion) geringer zeitlicher Abstände (DT_1) und, zur Kompensation der dafür notwendigen zusätzlichen Zeit, eine Kompression großer zeitlicher Abstände (DT_4) eingeführt (Bild 6c). Das ist eine Kompandierung der Zeitbasis des Signals, die am Empfänger durch eine inverse Dekompandierung wieder rückgängig gemacht werden kann.

Bild 7 zeigt die Umrechnung der Zeitdifferenzen DT in die kompandierten Zeitdifferenzen DTk über die Kompanderkennlinie KL. Für diese lineare Kennlinie gilt:

$$\mathrm{DTk} = \mathrm{DTk}_0 + k \cdot \mathrm{DT} \tag{3}$$

$$k = \frac{\mathrm{DTk\,max} - \mathrm{DTk\,min}}{\mathrm{DT\,max} - \mathrm{DT\,min}} = \frac{\mathrm{DTk\,max} - \mathrm{DTk}_0}{\mathrm{DT\,max}}\,. \tag{4}$$

Dabei kann $\mathrm{DT}_{\min}$ theoretisch sogar Null sein, nur nicht zu oft. Das Verfahren ist damit nicht auf bandbegrenzte Verfahren beschränkt. Seltene hochfrequente Signaländerungen, wie feinste Details im Bild, werden lupenhaft gedehnt, grobe Details werden komprimiert dargestellt (vor der Dekompandierung).

Aus Bild 7 wird deutlich, wie die Varianz der zeitlichen Differenz DT der Abtastungen durch die Spiegelung der DT-Achse an der Kompanderkennlinie bei $k < 1$ verringert wird. Daraus und aus Bild 6d wird deutlich, daß diese Kompandierung zu einer nichtlinearen Bandbreitenkompression führt. Die obere Frequenzgrenze wird geringer, die untere Grenzfrequenz steigt.

Das geht mit linearer Signalverarbeitung bekanntlich nicht. Die arhythmischen Abtastverfahren sind aber nichtlineare Verfahren. Sie haben unbestrittene Vorteile bei der redundanzreduzierten Übertragung und Speicherung von Signalen durch ihre signaladaptive Abtastung und lassen hier nun neue Möglichkeiten für die Signalbearbeitung erkennen.

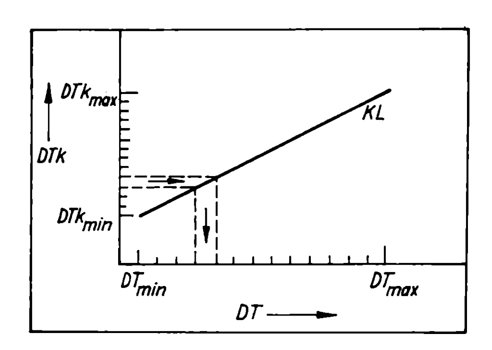


Bild 9. Zeitauflösungsänderung bei Kompandierung und Dekompandierung

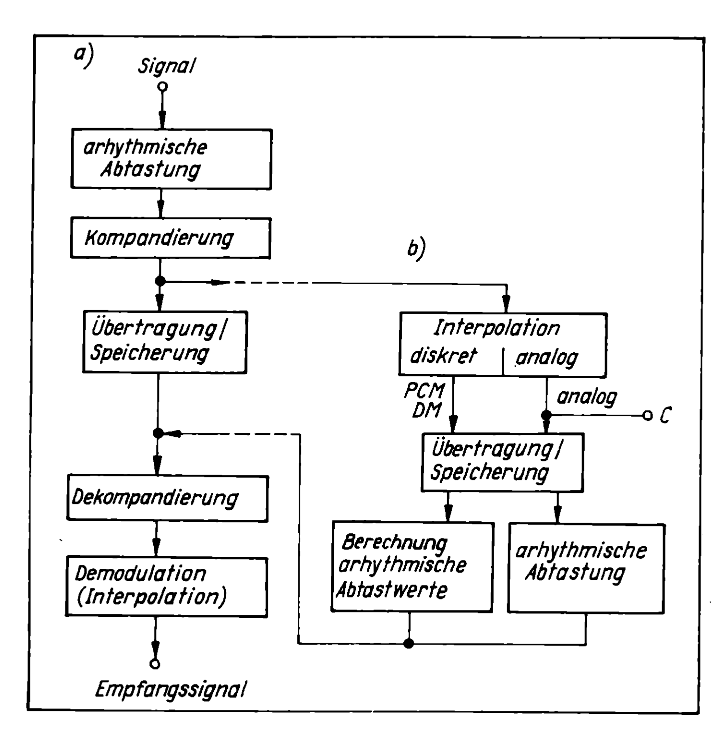


Bild 10. Blockschaltbild ableitbarer Verfahren zur zeitkompandierten Übertragung und Speicherung von Signalen
a) arhythmische Übertragung
b) analoge oder rhythmische Übertragung

Die arhythmische signaladaptive Abtastung der EWM ist ja nicht an die Bandbreite des Signals angepaßt, sondern folgt nur der Schwankung der oberen Grenzfrequenz. Die zeitliche Kompandierung, so scheint es nach dem gegenwärtigen Erkenntnisstand, führt zur Kompression der Bandbreite, wenn ein großer Teil der Bandbreite nur zeitweilig belegt ist und führt somit zur Anpassung an die momentane Bandbreite.

Bild 8 zeigt weitere mögliche Kompanderkennlinien. Sie können gekrümmt sein, brauchen nicht einmal stetig zu sein, müssen aber eineindeutig sein.

KL2 ist eine gekrümmte Kennlinie, die den unten erklärten Zeitauflösungsverlust mehr auf tiefe Frequenzen verschiebt.

KL3 ist die Kennlinie, mit der nur hohe und tiefe Spektralteile beeinflußt werden.

KL4 ist eine nichtstetige Kennlinie, mit der ein Bereich von DTk-Werten für die Signalübertragung ausgeschlossen wird. Diese DTk-Werte sind frei für Alarm, Signalisierung in Sprachpausen u.a.m.

KL5 ist eine Kennlinie zur Bandbreitenerhöhung, und KL6 ist eine Invertierungskennlinie, die z.B. der Sprachverschlüsselung dienen kann.

Komprimiert man ein Signal auf die halbe Bandbreite und überträgt es über einen verrauschten Kanal (mit weißem Rauschen), wird die Rauschleistung aufgrund der halbierten Bandbreite nur halb so groß sein wie bei der Übertragung in Originalbandbreite. Das heißt, man gewinnt für das kompandierte Signal theoretisch 3 dB (genauer: $-10 \lg k$ dB) im Signal-Rausch-Verhältnis. So weit gehen die Vorteile nun doch nicht! Dieser Gewinn muß wie Bild 9 zeigt, bei der Dekompandierung als zeitlicher Auflösungsverlust wiedergegeben werden.

Auf der vertikalen Achse ist im Bild 9 die am Empfänger mögliche, durch das Rauschen begrenzte, zeitliche Auflösung aufgetragen, die bei der Dekompandierung entsprechend dem Anstieg der Kennlinie auf der horizontalen Achse um den Faktor $1/k$ gedehnt wird. Durch gekrümmte Kennlinien (KL2 im Bild 8) kann der Auflösungsverlust ungleichmäßig verteilt werden. Zweckmäßigerweise wird er so verteilt, daß bei hohen Frequenzen die Lage der Abtastwerte genauer und bei tiefen Frequenzen ungenauer reproduziert wird, weil die Umrechnung von Lageabweichungen in Amplitudenabweichungen frequenzproportional ist.

3. Anwendung der Zeitkompandierung

3.1. *Übertragung und Speicherung zeitkompandierter Signale*

Die Übertragung und Speicherung von Ton- und Bildsignalen zur Reproduktion der Signale schließt die Dekompandierung vor der Demodulation im Empfänger ein. Die notwendigen Schritte dafür zeigt Bild 10.

Im Bild 10a ist die arhythmische zeitlich kompandierte Abtastfolge bzw. Impulsfolge zu übertragen und zu speichern. Mit zwei zusätzlichen Signalverarbeitungsschritten ist, nach Bild 10b, auch die Übertragung als zeitkompandiertes analoges Signal (nach Bild 6d) oder auch als rhythmische Abtastwertfolge (PCM) möglich.

3.2. *Zeitkompandierte Signaldarstellung*

Für die Bildmustererkennung und für die Auswertung von Meßsignalen sind die wenig Änderungen enthaltenen Signalabschnitte i.allg. weniger interessant, während feine Details gern zeitlupengedehnt beobachtet werden. Arhythmische signalabhängige Abtastverfahren sind naheliegend und die gewünschte Darstellung kann durch die zeitkompandierte Registrierung und Darstellung unterstützt werden. Dieses Signal ist im Bild 10 am Punkt C abnehmbar. Die Darstellung ist dann nichtlinear verzerrt, kann aber zur meßtechnischen Auswertung wieder durch Dekompandierung entzerrt werden.

4. Zusammenfassung

Die vorgestellte Zeitbasiskompandierung [9] ermöglicht eine bessere Ausnutzung der Übertragungszeit bei arhythmischer Codierung von Signalen, deren momentanes Spektrum starken Änderungen unterliegt und die Langzeitbandbreite stets nur zu einem geringen Teil ausfüllt. Im Zusammenhang mit der arhythmischen Abtastung ist es ein nichtlineares Verfahren, das neue Wege der Signalbearbeitung zur redundanzarmen Übertragung, Speicherung und Darstellung erschließt.

Dazu müssen noch einige Probleme, wie Grenzen der Kompandierung, geeignete Kennlinien und die Interpolation in Abhängigkeit vom Lang- und Kurzzeitspektrum des Signals untersucht werden.

NaA 31

Literatur

[1] *Zühlke, W.:* DA, DT-Modulation, ein signaladaptives Codierverfahren. 32. Intern. Wiss. Koll. TH Ilmenau 1987, H. 2, S. 101–106
[2] *Steele, R.:* Delta modulation systems. London: Pentech Press, 1975
[3] *Matthews, M. V.:* Extremal coding for speech transmission. IRE Trans. on. inf. Theory 5 (1959) 10, S. 129–136
[4] *Glaser, W.; König, D.:* Die Extremwertmodulation als Verfahren zur Verringerung der mittleren Problemrate bei der Übertragung von Sprache. Nachrichtentechnik 19 (1969) 10, S. 368–374
[5] *Zühlke, W.:* Untersuchungen zur Extremwertmodulation mit Sprachsignalen. Nachrichtentech., Elektron. 27 (1977) 2, S. 500–503
[6] *Späth, H.:* Spline-Algorithmen. München Wien: R. Oldenbourg, Verlag 1973
[7] *Bronstein-Semendjajew:* Taschenbuch der Mathematik. Leipzig: Teubner Verlagsgesellschaft 1962, S. 517 ff
[8] *Bribach, J.:* Datenreduktion an einer autonomen anwählbaren seismischen Registrierstation, noch unveröffentlichte Diss. an der AdW, ZIPE Potsdam
[9] *Zühlke, W.:* Verfahren zur Änderung der Zeitbasis eines Signals. WP H04B/3114582

Dr. sc. techn. *Werner Zühlke*, Technische Hochschule Ilmenau, Sektion Informationstechnik und theoretische Elektrotechnik, Postschließfach 327, Ilmenau, 6300

Präzisionswinkelmodulation für CPM-Verfahren

H. Hiller; C. Wobus, KDT, Dresden

Mitteilung aus der Technischen Universität Dresden, Sektion Informationstechnik

Die Anwendung verschiedenster Winkelmodulationsverfahren stand auf dem Gebiet der Modulationstheorie für die Übertragung von Binärfolgen in den letzten Jahren erneut im Mittelpunkt des Interesses. Eine Vielzahl von Untersuchungen wurde vorgenommen, um für den durch die zu übertragende Binärfolge modulierten Träger eine Signalgestaltung zu erzielen, für die die spektrale Leistungsdichte in einem schmalen Übertragungsband konzentriert ist. In gewissem Umfang wurde auf die historische Seite dieser Entwicklung in [1] eingegangen. Bezüglich der Signalgestaltung kann man feststellen, daß die Entwicklung dahin geht, durch eine Modulation des Winkels des Trägers einen geglätteten Verlauf zu erzielen, der zwischen PSK und FSK einzuordnen ist. Dabei erreichen die CPM-Verfahren (Continous Phase Modulation), die zu den Zeitpunkten der Bitgrenzen der modulierenden Binärfolge einen geglätteten Verlauf haben die kleinsten Bandbreiten. Weitere Möglichkeiten der Bandbreitereduktion bestehen durch Vorkorrelation in der Folge bzw. durch die Konstruktion mehrvalenter Verfahren.

Die in der Arbeit [2] durchgeführten Untersuchungen zeigen auch, daß die Berücksichtigung der Vor- und Nachgeschichte eines Binärsymbols der zu übertragenden Folge für die Gestaltung des Phasenverlaufs von Bedeutung ist.

Es gibt damit eine Reihe von Gesichtspunkten, deren Beachtung zu einem schmalbandigen Signalspektrum führen kann.

Sendeseitig bleibt jedoch zunächst das Hauptproblem, über einen Modulator zu verfügen, mit dessen Hilfe die Modulation des Trägersignals tatsächlich so präzise vorgenommen werden kann, daß die gewünschte spektrale Leistungsverteilung entsteht. Prinzipiell besteht zunächst die Möglichkeit, die Winkelmodulation mit einem spannungsgesteuerten Oszillator (VCO) zu realisieren. In der Praxis spricht die relative, geringe Frequenzkonstanz bisheriger Schaltungen gegen eine solche Lösung. Methoden der Trägerfrequenzregelung [2] [3] führen zu vermaschten mehrschleifigen Regelkreisen und damit zu notwendigen Einschwingzeiten, bis die Senderanordnung überhaupt „modulationsfähig" ist. Dabei wurde vor allem in Erwägung gezogen, daß für die Übertragung digitalisierter Sprache ein ständiger Wechsel zwischen Senden und Empfangen vollzogen wird.

Eine zweite Möglichkeit besteht in der Verwendung eines Phasenregelkreises, indem PSK-Signale durch diesen einer Phasenglättung bezüglich des Trägersignals unterzogen werden [4]. Problematisch ist dort wiederum eine solche Gestaltung des Schleifenfilters, daß die Phasenübertragungsfunktion einschließlich der PSK-Phasensprünge bestimmte Nyquistkriterien erfüllt.

Eine weitere bisher angewendete Modulatorstruktur ist im Tabellensuchverfahren unter Verwendung eines Quadraturmodulators zu sehen.

1. Tabellensuchverfahren mit Quadraturmodulator

Der Modulator (Bild 1) wird mit einer zu übertragenden Binärfolge gespeist. Da die Phasenverläufe konkret vorher berechnet wurden, müssen diese Verläufe für ein zu übertragendes Binärzeichen unter Berücksichtigung von $(R-1)/2$ Binärzeichen Vorgeschichte und $(R-1)/2$ Binärzeichen Nachgeschichte bei einer Schieberegisterlänge R gespeichert sein. Das bedeutet aber andererseits auch, daß für die Dauer T_B eines Binärschrittes 2^R verschiedene Phasenverläufe

$$\varphi_\nu(t);\ \nu = 0 \cdots 2^R - 1$$

$$(k-1)\,T_B \leqq t \leqq k \cdot T_B$$

verfügbar sein müssen. Das vom Modulator zu erhaltende Ausgangssignal hat dann die Form

$$u_a(\nu, t) = U_0 \sin[\omega_0 t + \varphi_\nu(t)]. \tag{1}$$

Da bisher die direkte Beeinflussung des Arguments der sin-Funktion mit Hilfe eines VCO ausschied, wurde eine Synthetisierung des modulierten Trägers entsprechend des Additionstheorems

$$\sin[\omega_0 t + \varphi_\nu(t)] = \sin\omega_0 t \cos\varphi_\nu(t) + \cos\omega_0 t \sin\varphi_\nu(t) \tag{2}$$

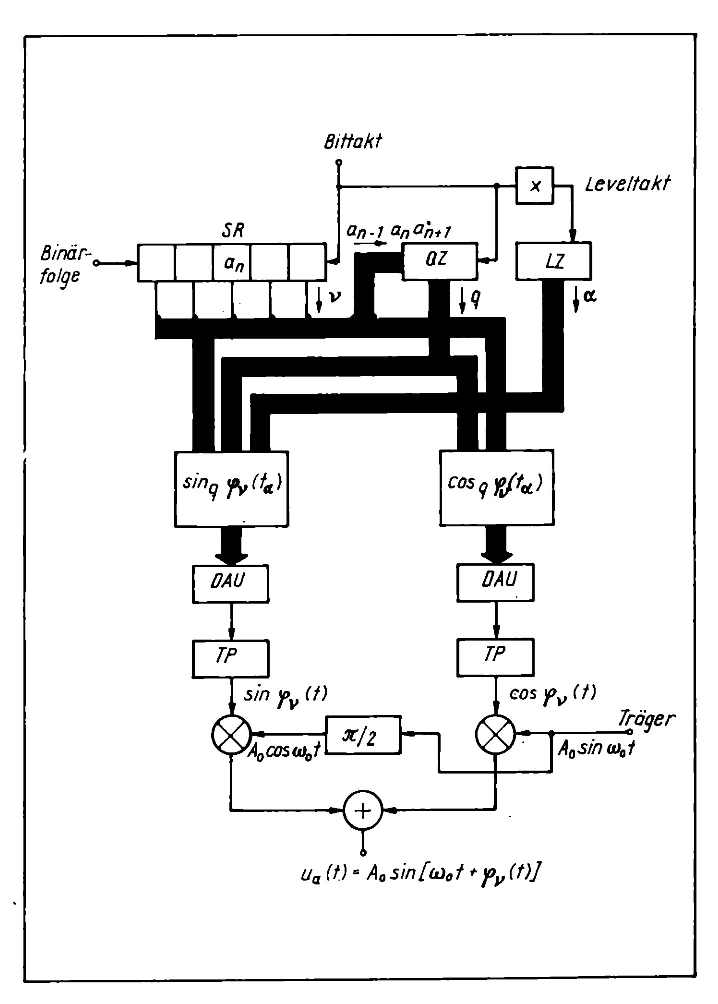
vorgenommen. Daraus folgt für den Modulator:

- Es müssen nicht die Phasentrajektorien $\varphi_\nu(t)$, sondern deren sin- bzw. cos-förmige Verläufe in den Tabellenspeichern abgelegt werden.
- Da der Tabellenspeicher nur eine endliche Anzahl diskreter Stützwerte enthalten kann, muß eine Zeitdiskretisierung $(t \to t_\alpha)$ durchgeführt werden.
- Um stetige Amplitudenanschlüsse zu gewährleisten, ist für den modulierten Trägerzeiger zu kontrollieren, welcher Quadrant erreicht wurde. Der Träger darf durch die Modulation während der Dauer eines Binärzeichens nur einen Phasenhub von 0, $\pm\pi/4$, $\pm\pi/2$ erfahren.

Die Schaltung nach Bild 1 berücksichtigt diese Forderungen. Eine Anzahl von R Symbolen der Folge bestimmt die auszuwählende Trajektorie $\varphi_\nu(t)$ bzw. deren Index ν. Die mittleren 3 Binärzeichen bestimmen den Phasenhub während der Dauer T_B eines Zeichens und steuern den Quadrantenzähler QZ, über den die richtigen Verläufe des sin bzw. cos und durch den Levelzähler LZ die Stützwerte zu den Zeitpunkten t ausgelassen werden. Nachfolgende Digital-Analog-Umsetzer, Tiefpaßfilter und Quadraturmodulatoren führen zum approximierten Verlauf der modulierten Schwingung nach (1) bzw. (2). Untersuchungen [5] haben jedoch gezeigt, daß es einige Realisierungsprobleme gibt:

- die analoge Phasendrehung des sinusförmigen Trägers um exakt 90°

Bild 1. CPM-Modulator mit Quadraturmodulator unter Anwendung des Tabellensuchverfahrens



- die Multiplikatoren müssen eine ausreichende Dynamik bei hohen Linearitätsforderungen erfüllen (mit B 222 oder K 140 MA 1 nicht gewährleistet)
- die Verwendung zweier Tabellenspeicher, zweier DAU und Tiefpaßfilter läßt den Aufwand ansteigen, wobei Bild 1 den Trägergenerator noch nicht mit darstellt
- eine effektive Digitalisierung des Modulationsvorgangs ist diese Schaltungsanordnung noch nicht.

Es wurde deshalb eine Schaltungsanordnung entwickelt, aufgebaut und untersucht [6], die inzwischen auch von anderen Autoren, unabhängig von diesen Untersuchungen, bestätigt wurde [7].

2. Digitaler Präzisionswinkelmodulator

Ausgangspunkt für einen digitalen Präzisionswinkelmodulator ist (1). Der Grundgedanke ist dabei, daß die von einem Quarzgenerator der Frequenz $1/T_I$ erhaltenen Impulse einem Adressenzähler zugeführt werden, der die Adressen A_α mit jedem Takt nach

$$A_\alpha: = [A_{\alpha-1} + N] \bmod K; \quad A_0: = 0 \tag{3}$$

neu bildet. Mit diesen Adressen werden die Speicherinhalte eines EPROM aufgerufen, der K binär dargestellte Stützwerte einer sinusförmigen Schwingung einer Periode (0 bis 2π) enthält. Das Ausgangssignal eines nachfolgenden Digital-Analog-Umsetzers hat dann zu den Zeitpunkten t_α die Stützwerte

$$u_a(t_\alpha) = U_0 \cdot \sin \frac{N}{T_I} \cdot \frac{2\pi}{K} \cdot t_\alpha \tag{4}$$

bzw. nach der TP-Filterung die Form

$$u_a(t) = U_0 \cdot \sin \frac{N}{T_I} \cdot \frac{2\pi}{K} \cdot t.$$

Die Modulation des Trägers kann nur dadurch vorgenommen werden, daß mit jeder Adressenbildung für das Auslesen aus dem sin-EPROM nicht um eine Zahl N, sondern um mehr oder weniger als N weitergezählt wird. Um diesen Gedanken auszuformulieren, muß jedoch auf die Phasentrajektorien $\varphi_\nu(t)$ nochmals eingegangen werden. Da nach der beschriebenen Methode die Phasenänderung entsprechend der Adreßänderung akkumulativ erfolgt, sind aus den Phasentrajektorien $\varphi_\nu(t)$ die Änderungen zwischen 2 diskreten Zeitpunkten t_α und $t_{\alpha-1}$ zu ermitteln. Die Änderungen sind damit nach

$$\frac{\Delta\varphi_\nu(t_\alpha)}{T_I} = \frac{\varphi_\nu(t_\alpha) - \varphi_\nu(t_{\alpha-1})}{t_\alpha - t_{\alpha-1}} \tag{5}$$

und die zusätzlichen in (3) zu berücksichtigenden Adressenänderungen nach

$$\Delta A_\nu(t_\alpha) = \frac{\Delta\varphi_\nu(t_\alpha)}{2\pi/K} \tag{6}$$

bzw.

$$\Delta A_\nu(t_\alpha) = \text{integer}\left\{\frac{\Delta\varphi_\nu(t_\alpha)}{2\pi/K}\right\}$$

zu berechnen. Die Erfüllung der Bedingung

$$\frac{2\pi}{K}\sum_{\mu=1}^{T_B/T_I} \Delta A_\nu(t_\mu) = 0;\ \pm\frac{\pi}{4};\ \pm\frac{\pi}{2}$$

ermöglicht es, die Akkumulierung von Quantisierungsfehlern zu vermeiden. Aus (3) wird nun

$$A_\alpha: = [A_{\alpha-1} + N + \Delta A_\nu(t_\alpha)] \bmod K$$

bzw. unter Einbeziehung eines konstanten Adressenoffsets AC

$$A_\alpha: = [A_{\alpha-1} + (N - AC) + AC + \Delta A_\nu(t_\alpha)] \bmod K \tag{7}$$

mit $A_0: = 0$ und $AC + \Delta A_\nu(t_\alpha) \geqq 0$. Die Stützwerte des modulierten Trägers werden durch

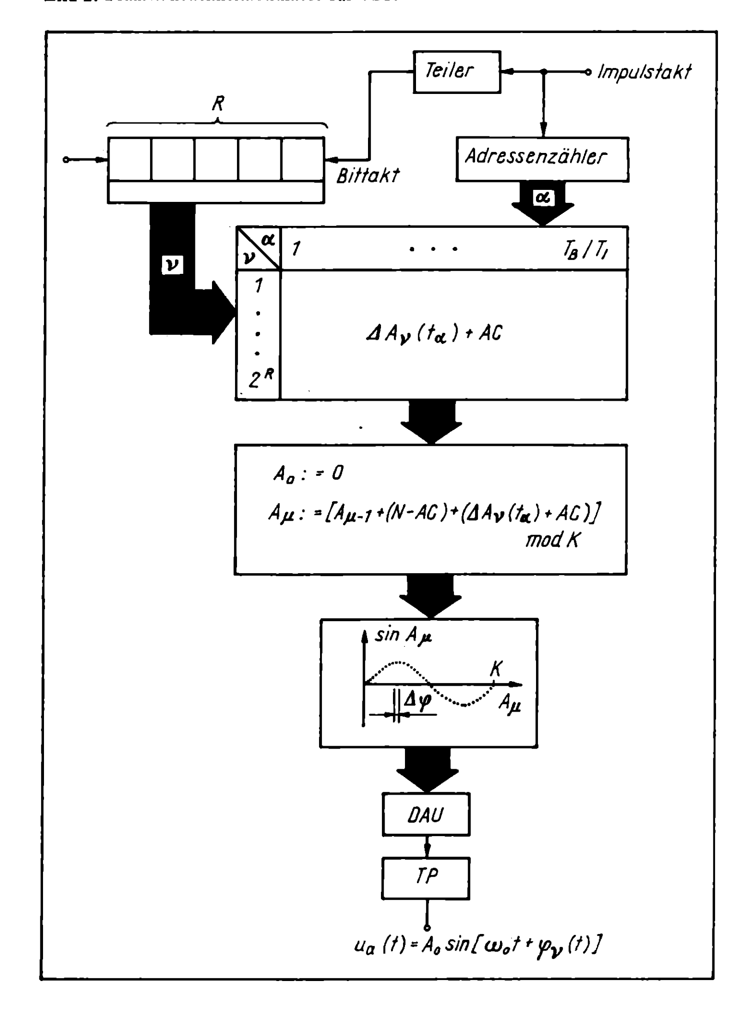
$$u_a(\nu, t_\alpha) = A_0 \sin\left[\frac{N}{T_I}\frac{2\pi}{K} t_\alpha + \sum_{\mu=1}^{\alpha} \Delta A_\nu(t_\mu)\frac{2\pi}{K}\right] \tag{8}$$

bzw. nach TP-Glättung ein phasenmoduliertes Signal

$u_a(\nu, t) = A_0 \sin[\omega_0 t + \varphi_\nu(t)]$ ergeben.

Die zugehörige Schaltungsanordnung ist im Bild 2 dargestellt.
Das Datensymbol a_n sowie $(R-1)/2$ Symbole Vorgeschichte $(R-1)/2$ Symbole Nachgeschichte bestimmen, welche der 2^R Phasentrajektorien den Träger für die Dauer dieses Binärschrittes modulieren wird; bestimmen den Index ν. Der Adressenzähler ruft zu den Zeitpunkten t_α die niederwertigen Adressenteile auf, wodurch die zu einer Phasentrajektorie $\varphi_\nu(t)$ gehörenden Adressenänderungen $\Delta A_\nu(t_\alpha) + AC$ der nachfolgenden Summationsschaltung zugeführt werden können. Der konstante Adressenoffset hat hierbei zum Ziel, nur mit positiven Speicherinhalten zu arbeiten und nachfolgend tatsächlich nur zu addieren. Damit ist bereits beschrieben, daß die nachfolgende Additionsschaltung die Adressenberechnung nach (7) realisiert. Die somit aus dem sin-EPROM ausgelesenen Stützwerte des modulierten Trägers präsentieren sich entsprechend (8) in der Form von Binärworten der Breite m am Ausgang des Digital-Analog-Umsetzers in amplitudendiskreter Form. Die gewünschte Form des modulierten Trägers entsteht nach einer TP-Filterung.
Die Funktion dieser Schaltung wird durch das Beispiel (Bild 3 und Tafel 1) veranschaulicht. Im Bild 3 wird im 1. Diagramm ein Teil der modulierten Folge gezeigt. Die darunter gezeigten Diagramme stellen die Vorgänge für die Dauer dieses einen Binärschritts in Zeitspreizung dar. Dies heißt zunächst, daß das Trägersignal bzw. das modulierte Trägersignal während dieser Zeit mit $T_B/T_I = 32$ Stützwerten bzw. zu 32 diskreten Zeitpunkten t_α synthetisiert werden soll. Es wird angenommen, im sin-EPROM seien 256 Amplitudenwerte gespeichert und die Additionsschaltung zählt die Adressen um $N = 25$ weiter.
Für die 1. Schwingung, die dargestellt wurde, gilt die Annahme, daß alle Adressenänderungen $\Delta A_\nu(t_\alpha) = 0$ sind. Damit wird die unmodulierte, synthesierte Trägerschwingung erhalten. Die sich ergebenden Adressen A_α nach (3) und die auf U_0 normierten Stützwerte entsprechend (4) sind in Tafel 1 aufgeführt. Desgleichen sind die Stützwerte und die interpolierte Trägerschwingung im Bild dargestellt.
Der nächste Schritt bezieht eine beispielhafte Phasentrajektorie $\varphi_\nu(t)$ ein, die durch eine Näherung $\varphi_\nu(t_\alpha)$ dargestellt wird. Die gesamte Phasendrehung soll nach der Zeitdauer T_B einen Wert von $\Delta\varphi = \pi/2$ betragen. Das bedeutet für eine Phasenänderung von 2π über 256 Adressen eine Adressenänderung um zusätzlich 64 Adressen während des Binärschrittes. Damit lassen sich aus der Kurve $\varphi_\nu(t_\alpha)$ die den $\Delta\varphi_\nu(t_\alpha)$ (5) entsprechenden Adressenänderungen $\Delta A_\nu(t_\alpha)$ ermitteln. Die zugehörigen diskreten normierten Amplituden nach (8) sind mit den Adressenänderungen in Tafel 1 nach-

Bild 2. Präzisionswinkelmodulator für CPM-Verfahren



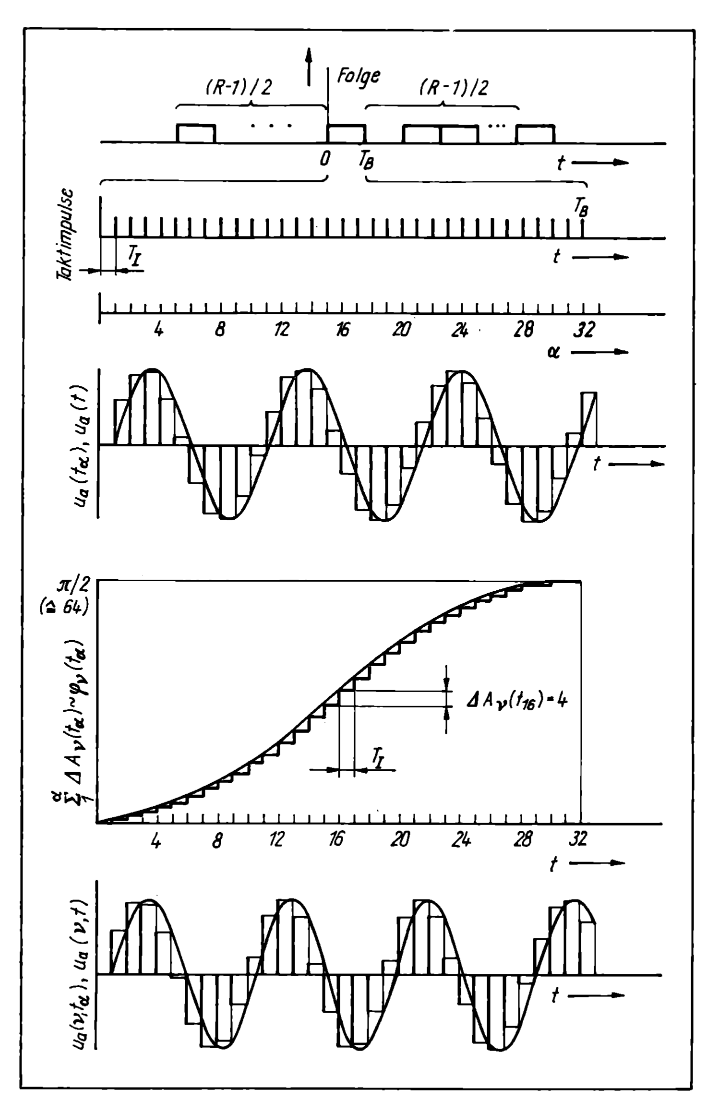


Bild 3. Beispiel einer Signalsynthetisierung durch einen Präzisionswinkelmodulator

vollziehbar. Die Stützwerte und die interpolierte modulierte Trägerschwingung sind im Bild 3 ebenfalls dargestellt und zeigen die erreichte Phasendrehung um $\pi/2$ gegenüber dem unmodulierten Träger. Bei dieser Beschreibung der Funktion der Schaltung wurde auf eine weitere detaillierte Darstellung der Taktung der einzelnen Baugruppen oder der Übernahme der binären Ausgangssignale in Latches verzichtet, um die Verständlichkeit nicht zu erschweren.

3. Abschätzung der Quantisierungsverzerrungen

Durch die digitale Signalgenerierung entstehen zwei voneinander unabhängige Quantisierungsfehler:

1. Ein dem modulierten Signal überlagertes Quantisierungsgeräusch aufgrund der endlichen Wortbreite m (z.B. $m = 8$ bit) im sin-EPROM und DA-Umsetzer.
2. Anstelle des Sollwerts $\varphi_\nu(t_\alpha)$ muß der nächstliegende der K gespeicherten Phasenwerte verwendet werden.

Der maximale Fehler im Zeitbereich (unter Vernachlässigung von Fehlern der DAU) entspricht bei dieser linearen Quantisierung jeweils der Hälfte der Stufenbreite $\left(2^{-(m-1)} \text{ bzw. } \frac{2}{K}\right)$. Durch das Glättungsfilter nach dem DA-Umsetzer kann der Fehler weiter reduziert werden.

Für den Modulatorentwurf ist der Einfluß auf das Spektrum interessant. Für das durch die Amplitudenquantisierung des Sinus bedingte Rauschen kann die aus der PCM-Technik bekannte Formel für das Signal-Quantisierungsgeräusch-Leistungsverhältnis

$$\frac{P_s}{P_{r\,\mathrm{ges}}} = 3 \cdot 2^{2m-1}$$

verwendet werden. Bei der Berechnung des Quantisierungsgeräusches in einem bestimmten Frequenzband (z.B. Nachbarkanal) ist die im Verhältnis zum Nutzsignal viel breitere spektrale Verteilung des Quantisierungsgeräuschs zu beachten. Für gestufte Rechtecksignale ist der Anteil des Quantisierungsgeräusch in einem Frequenzband

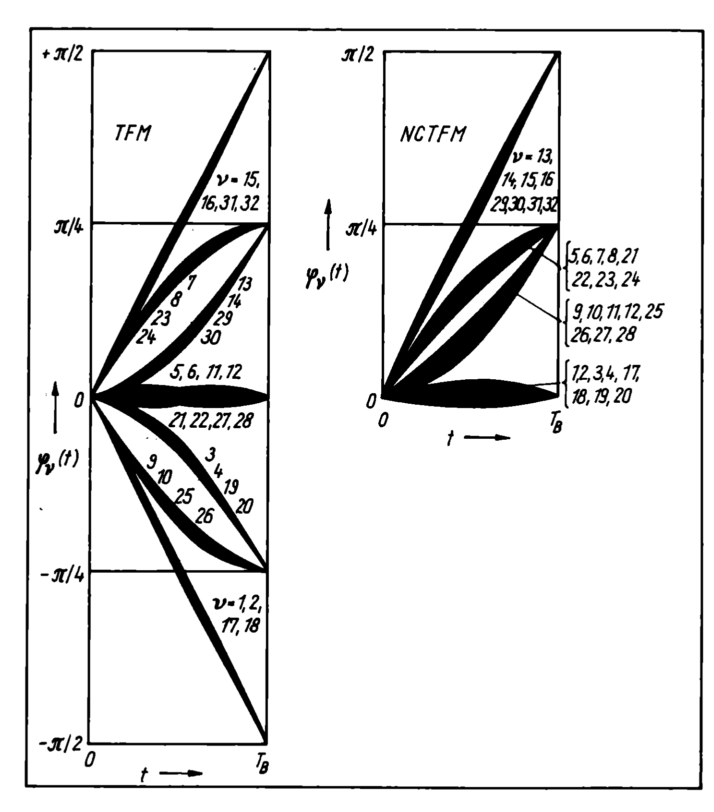
$$\frac{P_r}{P_{r\,\mathrm{ges}}} = 2\,T_\mathrm{I} \int_{f_1}^{f_2} sp^2(\pi f\,T_\mathrm{I})\,\mathrm{d}f.$$

Der Einfluß des Phasenfehlers ist schwerer abzuschätzen, da dieser Fehler das (nichtlineare) Winkelmodulationsverfahren durchläuft. In seiner relativen Größe ist er auf den Phasenhub im Bitintervall, z.B. $\frac{\pi}{2}$, zu beziehen.

Tafel 1. Zusammenstellung der zeitdiskreten Adressenänderungen und der damit entstehenden diskreten Trägeramplituden für den unmodulierten und den am Beispiel modulierten Träger

	A_μ	$u_\mathrm{T}(t_\alpha)/U_0$	$\Delta A_\nu(t_\alpha)$	A_μ	$u_a(t_\alpha)/U_0$
0	0	0.	0	0	0.
1	25	.5758	1	26	.5956
2	50	.9415	1	52	.9569
3	75	.9638	1	78	.9415
4	100	.5927	1	104	.5556
5	125	.0736	1	130	− .0491
6	150	− .5141	2	157	− .6532
7	175	− .9142	2	184	− .9808
8	200	− .9808	2	211	− .8932
9	225	− .6895	2	238	− .4276
10	250	− .1467	3	10	.243
11	19	.4496	2	37	.7884
12	44	.8819	3	65	.9997
13	69	.9925	3	93	.7572
14	94	.741	4	122	.1467
15	119	.2191	3	150	− .5141
16	144	− .3827	4	179	− .9495
17	169	− .8448	3	207	− .933
18	194	− .9988	4	236	− .4714
19	219	− .7883	3	8	.1951
20	244	− .2903	3	36	.7730
21	13	.3137	3	64	1.
22	38	.8032	2	91	.7884
23	63	.9997	2	118	.243
24	88	.8315	2	145	− .4052
25	113	.3599	2	172	− .8819
26	138	− .243	1	198	− .9892
27	163	− .7572	2	225	− .6895
28	188	− .9952	1	251	− .1224
29	213	− .8701	0	20	.4714
30	238	− .4276	1	46	.9039
31	7	.171	0	71	.9853
32	32	.7071		96	.7071

Bild 4. Phasentrajektorien für TFM und NCTFM



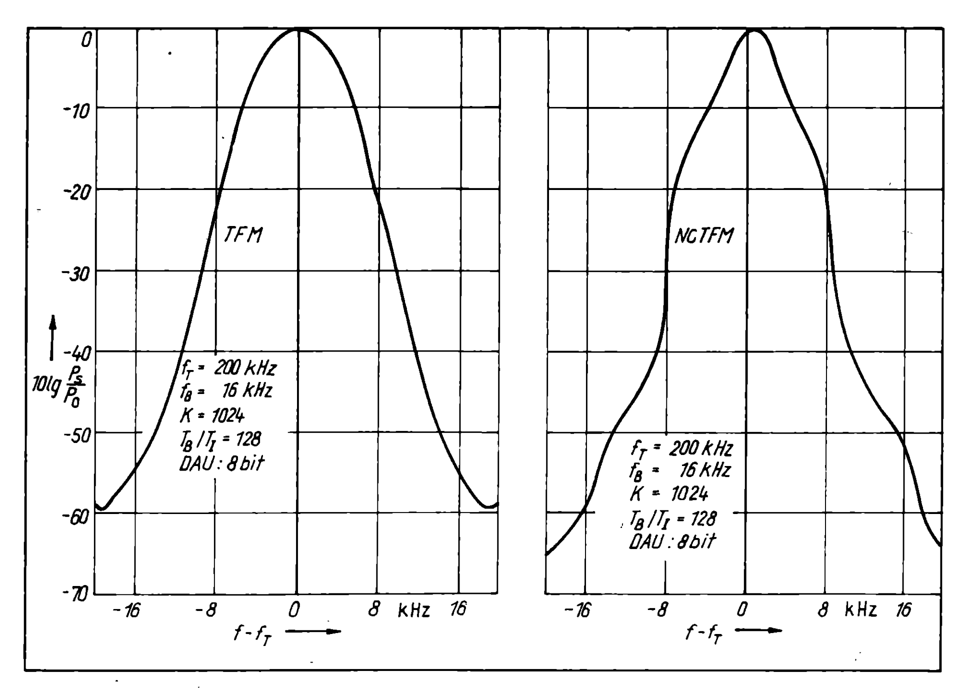


Bild 5. Spektrale Leistungsverteilung für TFM und NCTFM

Tafel 2. Zusammenstellung der wichtigsten Parameter des Präzisionsmodulators

Schieberegisterlänge	R
Anzahl d. Adr.-Leitungen d. Trajektorienspeichers	$R + ld\ T_B/T_I$
Anzahl der Stützwerte je Bit	T_B/T_I
Anzahl der Trajektorien	2^R
Folgentakt	$1/T_B$
Stützwerttakt	$1/T_I$
Anzahl der Stützwerte im Trajektorienspeicher	$2^R\ T_B/T_I$
Speichervolumen	2^R · Wortlänge · T_B/T_I
Speicherinhalt	$\Delta A_\nu (t_\alpha) + AC$
Adressenoffset	AC
Adressenänderung	$\Delta A_\nu (t_\alpha) \sim \Delta \varphi_\nu (t_\alpha)$
Anzahl der Stützwerte im sin-EPROM	K
Hub je Stützwert im sin-EPROM	$\Delta \varphi_{min} = \frac{2\pi}{K}$
Adressenänderung bei unmoduliertem Träger	N
Trägerfrequenz	$f_0 = \frac{N}{K} \cdot \frac{1}{T_I}$
maximaler Phasenhub (je T_I)	$\lvert \Delta A_{max} \rvert \cdot 2\pi/K$
Wortbreite des DAU	m

Der relative Quantisierungsfehler beträgt also maximal

$$\Delta x = \frac{2\pi}{K}.$$

Das entspricht einem dem Modulationssignal überlagerten Quantisierungsrauschen mit einem Abstand von $\frac{4}{3K^2}$.

Die Anzahl der Sinusabtastwerte K ist entsprechend groß zu wählen. Andererseits bereitet es in der analogen Schaltungstechnik einigen Aufwand, Signale mit solch großen Störabständen zur Verfügung zu stellen.

4. Erfahrungen

Die vorgestellte Konzeption für einen digital arbeitenden Präzisionsmodulator wurde an der TU Dresden auf der Grundlage des in der DDR 1986 verfügbaren Schaltkreissortiments als Funktionsmuster aufgebaut und erprobt [6]. Für die Erprobung in diesem Modulator wurden numerische Ergebnisse der Arbeit [2] für die Speicherung entsprechend Abschnitt 2 bearbeitet. Konkret betrifft das die CPM-Verfahren TFM und NCTFM. Die Ensemble der Trajektorien zeigt Bild 4, wobei auf die nähere Darstellung hier nicht eingegangen wird. In beiden Fällen beträgt die Folgenlänge R, die den Trajektorienverlauf bestimmt, $R = 5$. Damit enthält jeder Trajektorienspeicher 32 Verläufe. Die Bitfolgefrequenz beträgt $f_B = 16$ kHz, die Trägerfrequenz (als Zwischenfrequenz) $f_T = 200$ kHz, der sin-EPROM enthält 1024 Amplitudenwerte, und der Träger wurde während der Dauer eines Binärschritts aus $T_B/T_I = 128$ Stützwerten synthetisiert. Die experimentell gemessenen Spektren zeigten eine sehr gute Übereinstimmung mit den theoretisch berechneten Ergebnissen (Bild 5). Es bestätigte sich die Überlegung, daß die Amplitudenquantisierung und die Phasenquantisierung im wesentlichen das Quantisierungsrauschen im Übertragungskanal und in den Nachbarkanälen bestimmen.

Bei der Erzeugung von Bandpaß-Schmalbandsignalen kann auch mit Abtastfrequenzen unterhalb der Trägerfrequenz gearbeitet werden. Der erzielbaren Verminderung des Aufwands (Speicher, DA-Umsetzungsgeschwindigkeit) steht die Problematik der Trennung von Nutz- und Störspektrum gegenüber.

5. Zusammenfassung

Die Schaltungskonzeption des vorgestellten digitalen Präzisionsmodulators ermöglicht einen einfachen und übersichtlichen Entwurf eines CPM-Modulators, wobei die wichtigsten Parameter in Tafel 2 nochmals zusammengefaßt wurden. Es ist vorteilhaft, im experimentellen Betrieb durch den Wechsel des Trajektorienspeichers, das CPM-Verfahren ändern zu können.

Die digitale bzw. zeitdiskrete Arbeitsweise ermöglicht es weiterhin

- die zu erwartenden Fehler infolge Quantisierungsrauschen abzuschätzen
- höherentwickelte Schaltkreisstrukturen zu verwenden
- bzw. eine Schaltung zu integrieren.

Neben diesen grundlegenden Vorteilen gibt es für Teile der Schaltung weitere Anwendungsmöglichkeiten. Der Teil der Additionsschaltung und des sin-EPROM ist praktisch ein NCO (numerisch gesteuerten Oszillator) und das digitalisierte Gegenstück zum analog arbeitenden VCO. Der Einsatz in diskret arbeitenden digitalen Phasenregelkreisen ist damit möglich. Desgleichen ist es möglich, mit einem beliebigen zeit- und amplitudendiskreten Modulationssignal eine Präzisions-FM zu realisieren.

Der Teil des adressengesteuerten sin-EPROM, dessen Inhalt nicht nur sinusförmig sein muß, kann in Empfängerkreisen mit Entscheidungsrückführung [8] [9] als NCP (numerisch gesteuerter Phasenschieber) angewendet werden.

Die Anordnung des quarzgesteuerten Oszillators nach (4) ermöglicht es, auf diesem Weg phasenrauscharme Oszillatoren zu konzipieren. Allerdings können die Parameter und Frequenzen gegenwärtig ausgereifter Analoglösungen [10] [11] noch nicht erreicht werden.

Bezüglich der Realisierung eines Modulators sind bereits Entwicklungen durchgeführt worden [12] und es ist anzunehmen, daß sich diese Variante aufgrund ihrer Einfachheit und der vorliegenden Vorteile bei der Einführung in der Funktechnik durchsetzen könnte.

NaA 23

Literatur

[1] *Gieseler, S.; Engler, E.:* Modulationsverfahren mit kontinuierlichem Phasenverlauf zur Übertragung digitaler Signale. Nachrichtentech., Elektron. 35 (1985) 2, S. 67–69

[2] *Gieseler, S.:* Bandbreiteeffektive Modulationsverfahren für die Übertragung digitaler Informationen. TU Dresden, Dissertation A 1986

[3] *de Jager, F.; Dekker, C. B.; Muilwijk, D.:* System zur Datenübertragung mit Hilfe eines winkelmodulierten Trägers konstanter Amplitude. DE-OS 2838984, 15. 3. 79

[4] *Honma, K.; Murata, E.; Rikou, Y.:* On a Method of Constant Envelope Modulation for Digital Mobile Radio Communication. IEEE 1980, International Communication. S. 24. 1. 1.–24..1.5.

[5] *Krüger, J.:* TFM-Modulator. TU Dresden, Großer Beleg 1986

[6] *Krüger, J.:* Schmalband-Winkelmodulator. TU Dresden, Diplomarbeit 1987

[7] *Kopta, A.; Budisin, S.; Jovanovic, V.:* New Universal All-Digital CPM Modulator. IEEE Trans. on Commun. COM-35. (1987) 4, S. 458–462

[8] *Hiller, H.:* Eine verallgemeinerte Theorie zum Entwurf von Empfängern mit Entscheidungsrückführung für den Gaußschen Weißen Kanal, Wiss. Z. TU-Dresden 35 (1986) 5, S. 33–37

[9] *Hiller, H.:* Signalentscheidung und Synchronisation bei der Übertragung von diskreten, statistisch unabhängigen Nachrichten. TU Dresden, Dissertation B 1987

[10] *Jäckel, K.; Neumann, A.:* Phasenrauschanalyse von hochstabilen VHF-Quarzoszillatoren. Nachrichtentech., Elektron., 35 (1985) 2, S. 70–71

[11] *Nuszkowski, H.:* Baugruppe zur Erzeugung eines kurzzeitstabilen VHF-Signals. Techn. Mitteilungen des RFZ. 31. (1987) 3. S. 69–71

[12] *Suzuki, H.; Momma, K.; Yamao, Y.:* GMSK Digital Portable Transceiver Using 32 kbps ADM. IEEE 1983 Vehicular Technology Conference. S. 431–346

Dr. sc. techn. *Hartmut Hiller;* Dr.-Ing. *Christian Wobus*, Technische Universität Dresden, WB Kommunikations- und Computertechnik, Mommsenstr. 13, Dresden, 8027

Korrelations- und Spektralanalyse digitaler Signale

H.-J. Thomanek, KDT, Mittweida

Mitteilung aus der Ingenieurhochschule Mittweida, Sektion Informationselektronik

1. Problemstellung

Für die Dimensionierung von Systemen der digitalen Datenübertragung, beispielsweise für die Anpassung des Sendesignals an das Übertragungsmedium durch geeignete Leitungscodes oder auch im Zusammenhang mit der Extraktion der Taktinformation aus dem Nutzsignal, ist die Kenntnis und zielgerichtete Beeinflussung gewisser statistischer Eigenschaften der zufälligen digitalen Sendesignale nötig.

Eine exponierte Rolle spielt dabei die Ermittlung von Autokorrelationsfunktion und Leistungsdichtespektrum, eine Grundaufgabe der Signaltheorie, die in der Literatur (u.a. [1] bis [4]) breiten Raum einnimmt und nachfolgend für die Klasse der stochastischen Impulsprozesse (mit determiniertem Zeitraster) als allgemeines Modell zufälliger digitaler Signale behandelt wird.

Unter einem stochastischen Impulsprozeß wird im weiteren die zufällige Zeitfunktion

$$\xi(t) = \sum_{i=-\infty}^{\infty} \xi_i\, x(t - iT) \tag{1}$$

verstanden, darin sind $x(t)$ eine beschränkte und integrierbare determinierte Funktion und ξ_i eine stationäre Folge von Zufallsgrößen. Die Menge der ξ_i sei abzählbar und beschränkt.

Mit der Dirac-Funktion $\delta(t)$ folgt mit

$$\xi(t) = x(t) * \sum_{i=-\infty}^{\infty} \xi_i\, \delta(t - iT) = x(t) * \Delta(t) \tag{2}$$

die Darstellung des stochastischen Impulsprozesses durch eine Formfunktion $x(t)$ und eine Strukturfunktion $\Delta(t)$.

Für den Spezialfall $x(t) \equiv 0$ für $t < 0$ und $t > T$ zeigt Bild 1 eine mögliche Realisierung des stochastischen Impulsprozesses, der auch als periodischer oder Prozeß mit zyklischer Stationarität bezeichnet wird [5]. Prozesse dieser Art sind u.a. für die Dimensionierung von Leitungscodes in der digitalen Datenübertragung von Bedeutung.

2. Berechnung von Autokorrelationsfunktion und Leistungsdichtespektrum

Da $\Delta(t)$ und $\xi(t)$ Realisierungen (zyklisch) stationärer Prozesse darstellen, erhält man unter Berücksichtigung von (2) die Autokorrelationsfunktion des Impulsprozesses als Faltungsprodukt [6]

$$R_\xi(\tau) = R_x(\tau) * R_\Delta(\tau) \tag{3}$$

der Autokorrelationsfunktion der Formfunktion

$$R_x(\tau) = \int_{-\infty}^{\infty} x(t)\, x(t + \tau)\, \mathrm{d}t \tag{4}$$

und der Autokorrelationsfunktion der Strukturfunktion

$$R_\Delta(\tau) = \frac{1}{T} \sum_{n=-\infty}^{\infty} R(n)\, \delta(\tau - nT), \tag{5}$$

wobei letztere eine äquidistante Stoßfolge bildet. Die Koeffizienten $R(n)$ bezeichnen die Werte der Autokorrelationsfolge

$$R(n) = E\{\xi_i\, \xi_{i+n}\}. \tag{6}$$

Es gilt

$$R(n) = R(-n)$$
$$R(0) = E\{\xi_i^2\} = m_2$$
$$\lim_{n \to \infty} R(n) = [E\{\xi_i\}]^2 = m_1^2. \tag{7}$$

Mit den angeführten Beziehungen folgt aus (3) eine Darstellung der Autokorrelationsfunktion

$$R_\xi(\tau) = R_{\xi k}(\tau) + R_{\xi d}(\tau) \tag{8}$$

in einen konvergierenden

$$R_{\xi k}(\tau) = \frac{1}{T} \sum_{n=-\infty}^{\infty} B(n) \cdot R_x(\tau - nT) \tag{9}$$

und divergierenden Anteil

$$R_{\xi d}(\tau) = \frac{m_1^2}{T} \sum_{n=-\infty}^{\infty} R_x(\tau - nT). \tag{10}$$

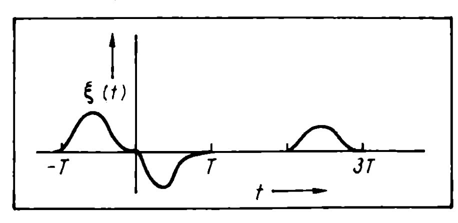
Zur vereinfachten Schreibweise wurde die Autokovarianzfolge

$$B(n) = E\{(\xi_i - E\{\xi_i\})(\xi_{i+n} - E\{\xi_{i+n}\})\}$$
$$= R(n) - m_1^2 \tag{11}$$

Tafel 1. Mittelwert und Streuung der Zufallsgröße ξ

Zufallsgröße	Lin. Mittelwert	Streuung
$\xi = +1, 0$ $(p = P\{\xi = +1\})$	p	$p(1 - p)$
$\xi = +1, -1$ $(p = P\{\xi = +1\})$	$2p - 1$	$4p(1 - p)$
$\xi_l = (l - 1);\ l = 1, 2, \ldots, m$ (gleichwahrscheinlich)	$(m - 1)/2$	$(m^2 - 1)/12$
$\xi_l = l - \frac{m+1}{2}$; $L = 1, 2, \ldots, m$ (gleichwahrscheinlich)	0	$(m^2 - 1)/12$

Bild 1. Realisierung eines stochastischen Impulsprozesses nach (1)



Tafel 2. Autokorrelationsfunktion und Leistungsdichtespektrum elementarer Impulssignale

Impulsform	Autokorrelationsfunktion	Leistungsdichtespektrum
Rechteckimpuls $x_0;\ \lvert t\rvert \leqq \frac{t_i}{2}$	$x_0^2 t_i \left(1 - \frac{\lvert\tau\rvert}{t_i}\right);\ \lvert\tau\rvert \leqq t_i$	$(x_0 t_i)^2 \cdot \mathrm{si}^2(\pi f t_i)$
Dreieckimpuls $x_0 \left(1 - 2\frac{\lvert t\rvert}{t_i}\right);\ \lvert t\rvert \leqq \frac{t_i}{2}$	$\frac{x_0^2 t_i}{3}\left[1 - 6\left(\frac{\tau}{t_i}\right)^2 + 6\left(\frac{\lvert\tau\rvert}{t_i}\right)^3\right];\ \lvert\tau\rvert \leqq \frac{t_i}{2}$ $\frac{x_0^2 t_i}{3}\left[2 - 6\frac{\lvert\tau\rvert}{t_i} + 6\left(\frac{\tau}{t_i}\right)^2 - 2\left(\frac{\lvert\tau\rvert}{t_i}\right)^3\right];\ \frac{t_i}{2} < \lvert\tau\rvert \leqq t_i$	$\left(\frac{x_0 t_i}{2}\right)^2 \cdot \mathrm{si}^4\left(\frac{\pi f t_i}{2}\right)$
cos-Impuls $x_0 \cos\left(\pi \frac{t}{t_i}\right);\ \lvert t\rvert \leqq \frac{t_i}{2}$	$\frac{x_0^2 t_i}{2}\left[\left(1 - \frac{\lvert\tau\rvert}{t_i}\right)\cos\left(\pi\frac{\tau}{t_i}\right) + \frac{\sin\left(\pi\frac{\lvert\tau\rvert}{t_i}\right)}{\pi}\right];\ \lvert\tau\rvert \leqq t_i$	$\left(\frac{2 x_0 t_i}{\pi}\right)^2 \cdot \frac{\cos^2(\pi f t_i)}{[1 - (2 f t_i)^2]^2}$
$\cos^2$-Impuls $x_0 \cos^2\left(\pi \frac{t}{t_i}\right);\ \lvert t\rvert \leqq \frac{t_i}{2}$	$\frac{x_0^2 t_i}{2}\left[\left(1 - \frac{\lvert\tau\rvert}{t_i}\right)\left(1 + \frac{\cos\left(2\pi\frac{\tau}{t_i}\right)}{2}\right) + \frac{3\sin\left(2\pi\frac{\lvert\tau\rvert}{t_i}\right)}{4\pi}\right];\ \lvert\tau\rvert \leqq t_i$	$\left(\frac{x_0 t_i}{2}\right)^2 \cdot \frac{\mathrm{si}^2(\pi f t_i)}{[1 - (f t_i)^2]^2}$
Gaußimpuls $x_0 \exp\left\{-\frac{t^2}{2a^2}\right\}$	$x_0^2 a \sqrt{\pi} \exp\left\{-\left(\frac{\tau}{2a}\right)^2\right\}$	$(a x_0)^2\, 2\pi$ $\exp\{-(2\pi f a)^2\}$

eingeführt. Fouriertransformation führt zum Leistungsdichtespektrum

$$S_\xi (f) = S_{\xi k} (f) + S_{\xi d} (f), \tag{12}$$

bestehend aus kontinuierlichem

$$S_{\xi k} (f) = \frac{1}{T} |X (f)|^2 \left[B (0) + 2 \sum_{n=1}^{\infty} B (n) \cos (2 \pi n f T) \right] \tag{13}$$

und diskretem Anteil

$$S_{\xi d} (f) = \frac{m_1^2}{T^2} \sum_{n=-\infty}^{\infty} \left| X \left(\frac{n}{T} \right) \right|^2 \delta \left(f - \frac{n}{T} \right). \tag{14}$$

Man erkennt, daß der Verlauf der spektralen Leistungsdichte durch die spektrale Amplitudendichte der Formfunktion

$$X (f) = \int_{-\infty}^{\infty} x (t) \cdot \exp \{-j 2 \pi f t\} \, dt \tag{15}$$

und mit der Autokovarianzfolge durch die Struktur (Codierung) der äquidistanten Stoßfolge $\Delta (t)$ bestimmt wird. Bemerkenswert ist die Invarianz der diskreten Komponente bezüglich der Struktur der äquidistanten Stoßfolge $\Delta (t)$.
Das eigentliche Problem besteht in der Ermittlung der Autokorrelations- bzw. Autokovarianzfolge. Trivial ist diese Aufgabe für den im Abschnitt 2.1 behandelten Spezialfall statistisch voneinander unabhängiger Elemente der Strukturfunktion $\Delta (t)$. Im Fall statistisch voneinander abhängiger Elemente ist die Bestimmung dieser Folgen unter Umständen problematisch. Geschlossen lösbar ist die in Abschnitt 2.2 dargestellte Modellierung der statistischen Bindungen durch Markov-Ketten. Im Zusammenhang mit der Untersuchung von Codes für die digitale Datenübertragung existieren Lösungen bei bestimmten Codierungsvorschriften (vgl. u. a. [6] [7]).

2.1. Impulsprozesse mit statistisch unabhängigen Impulsen

Bilden in (1) die Variablen ξ_i eine stationäre Folge unabhängiger Zufallsgrößen, so folgt mit

$$B (n) = \begin{cases} \sigma^2, & n = 0 \\ 0, & n \neq 0 \end{cases} \tag{16}$$

für den stochastischen Impulsprozeß die Autokorrelationsfunktion

$$R_\xi (\tau) = R_{\xi k} (\tau) + R_{\xi d} (\tau) = \frac{\sigma^2}{T} R_x (\tau) + \frac{m_1^2}{T} \sum_{n=-\infty}^{\infty} R_x (\tau - n T) \tag{17}$$

und das Leistungsdichtespektrum

$$S_\xi (f) = S_{\xi k} (f) + S_{\xi d} (f)$$

$$= \frac{\sigma^2}{T} |X (f)|^2 + \frac{m_1^2}{T^2} \sum_{n=-\infty}^{\infty} \left| X \left(\frac{n}{T} \right) \right|^2 \delta \left(f - \frac{n}{T} \right). \tag{18}$$

Streuung σ^2 und linearer Mittelwert m_1 der Zufallsvariablen ξ_i sind für einige interessierende Fälle in Tafel 1 angeführt. Tafel 2 enthält für bekannte Elementarimpulse die zugehörige Autokorrelationsfunktion $R_x (\tau)$ und das entsprechende Leistungsdichtespektrum $|X (f)|^2$.
Für den rechteckförmigen Elementarimpuls der Amplitude x_0 und der Breite $t_i = T$ erhält man den bekannten dreieckförmigen Verlauf der Autokorrelationsfunktion

$$R_{\xi k} (\tau) = x_0^2 \sigma^2 \left(1 - \frac{|\tau|}{T} \right); \quad |\tau| \leqq T$$

$$R_{\xi d} (\tau) = x_0^2 m_1^2 \tag{19}$$

und die durch die Spaltfunktion bestimmte spektrale Leistungsdichte

$$S_\xi (f) = x_0^2 \sigma^2 T \cdot \mathrm{si}^2 (\pi f T) + x_0^2 m_1^2 \sigma (f), \tag{20}$$

worin für den Spezialfall $m_1 = 0$ die den Gleichanteil repräsentierende diskrete Komponente verschwindet.

2.2. Markovsche Impulsprozesse

Unter Voraussetzung, daß die statistischen Bindungen der Werte ξ_i durch eine stationäre Markov-Kette 1. Ordnung mit insgesamt m Zuständen $\xi_l (i)$, $l = 1, 2, \ldots, m$ beschreibbar sind, berechnet sich die Autokorrelationsfolge zu

$$R (n) = \sum_{l=1}^{m} \sum_{k=1}^{m} \xi_l (i) \, \xi_k (i + n) \, q_l \cdot p_{lk} (n), \quad n \geqq 0. \tag{21}$$

Die Übergangswahrscheinlichkeiten $p_{lk} (n)$ der Übergangsmatrix $\mathbf{P}_n$ nach n Schritten erhält man mit

$$\mathbf{P}_n = \mathbf{P}_1^n \tag{22}$$

aus der gegebenen einschrittigen Übergangsmatrix $\mathbf{P}_1$ [8]. Die stationären totalen Wahrscheinlichkeiten q_l der Zustände $\xi_l (i)$ werden mit der Relation

$$p_l = \lim_{n \to \infty} p_{il} (n) = q_l \quad i, l = 1, 2, \ldots, m \tag{23}$$

aus den ergodischen Grenzwahrscheinlichkeiten p_l, deren Existenz im weiteren vorausgesetzt wird, berechnet [8].
Zur Ermittlung der n-ten Potenz der Matrix $\mathbf{P}_1$ verwendet man die bekannten Sätze aus der Theorie der Eigenwerte und Eigenvektoren reeller Matrizen. Letztendlich gilt für die transponierte Matrix (z. Beweis vgl. [3])

$$\mathbf{P}_n' = \mathbf{T} \, \underline{\Lambda}^n \, \mathbf{T}^{-1} \tag{24}$$

mit der potenzierten Eigenwertmatrix

$$\underline{\Lambda}^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdot & \cdot \\ \vdots & & \cdot & \cdot \\ 0 & \cdot & \cdot & \lambda_m^n \end{pmatrix} = \mathrm{diag} \left(\lambda_1^n, \lambda_2^n, \ldots, \lambda_m^n \right) \tag{25}$$

und der Matrix

$$\mathbf{T} = \begin{pmatrix} C_{11}^{(1)} & C_{11}^{(2)} & \cdots & C_{11}^{(m)} \\ C_{12}^{(1)} & C_{12}^{(2)} & \cdots & C_{12}^{(m)} \\ \vdots & \vdots & & \vdots \\ C_{1m}^{(1)} & C_{1m}^{(2)} & \cdots & C_{1m}^{(m)} \end{pmatrix}. \tag{26}$$

Die Eigenwerte λ_l mit $l = 1, 2, \ldots, m$ findet man aus den Wurzeln der algebraischen Gleichung m-ten Grades

$$\det (\mathbf{P}_1' - \lambda \mathbf{E}) = 0. \tag{27}$$

Hierin bezeichnet $\mathbf{E}$ die Einheitsmatrix. Die in (23) formulierte Existenz aller ergodischen Grenzwahrscheinlichkeiten p_l setzt voraus, daß ein Eigenwert $\lambda_{l_1} = 1$ ist und alle anderen Eigenwerte dem Betrag nach < 1 sind [3]. Vereinbart man den zum Eigenwert 1 gehörigen Index $l_1 = 1$, so gilt

$$\lim_{n \to \infty} \underline{\Lambda}^n = \mathrm{diag} (1, 0, \ldots, 0). \tag{28}$$

Die Elemente $C_{1i}^{(l)}$ mit $i, l = 1, 2, \ldots, m$ der Matrix $\mathbf{T}$ sind die aus der charakteristischen Matrix

$$\mathbf{C}^{(l)} = (\mathbf{P}_1' - \lambda_l \mathbf{E}), \; l = 1, 2, \ldots, m \tag{29}$$

berechneten Adjunkte. Verwendet man die in [3] eingeführte Matrizenschreibweise, so erhält man aus (21) die Autokorrelationsfolge

$$R (n) = \mathbf{e}' \mathbf{Q} \, \xi \, \mathbf{P}_n \, \xi \, \mathbf{e} = \mathbf{e}' \mathbf{Q} \, \xi \, (\mathbf{T}^{-1})' \, \underline{\Lambda}^n \, \mathbf{T}' \, \xi \, \mathbf{e}, \quad n \geqq 0 \tag{30}$$

mit den Diagonalmatrizen

$$\mathbf{Q} = \mathrm{diag} (q_1, q_2, \ldots, q_m) \tag{31}$$

$$\xi = \mathrm{diag} (\xi_1, \xi_2, \ldots, \xi_m) \tag{32}$$

und dem m-dimensionalen Vektor $\mathbf{e}' = (1, 1, \ldots, 1)$. Unter Beachtung von (7) und (28) errechnet sich die Autokovarianzfolge aus

$$B (n) = \mathbf{e}' \mathbf{Q} \, \xi \, (\mathbf{T}^{-1})' \, \underline{\Lambda}^n \, \mathbf{T}' \, \xi \, \mathbf{e}, \; \lambda_{l_1}^n = 0, \; n \geqq 0. \tag{33}$$

Folglich gelangt man von der Autokorrelationsfolge $R (n)$ zur Autokovarianzfolge $B (n)$, indem der Eigenwert $\lambda_{l_1}^n$ der potenzierten Eigenwertmatrix gleich Null gesetzt wird. Mit der Gleichung

$$m_1^2 = \lim_{n \to \infty} R (n) = \left[\sum_{l=1}^{m} p_l \cdot \xi_l \right]^2 \tag{34}$$

findet man den zur Auswertung von (10) und (14) ebenfalls notwendigen Mittelwert m_1. Der Einfluß der Formfunktion ist Tafel 2 zu entnehmen.

3. Spezielles Beispiel: binärer Markovscher Impulsprozeß

Zur Veranschaulichung und Weiterführung obiger Betrachtungen wird eine binäre Markovsche Impulsfolge mit der Übergangsmatrix

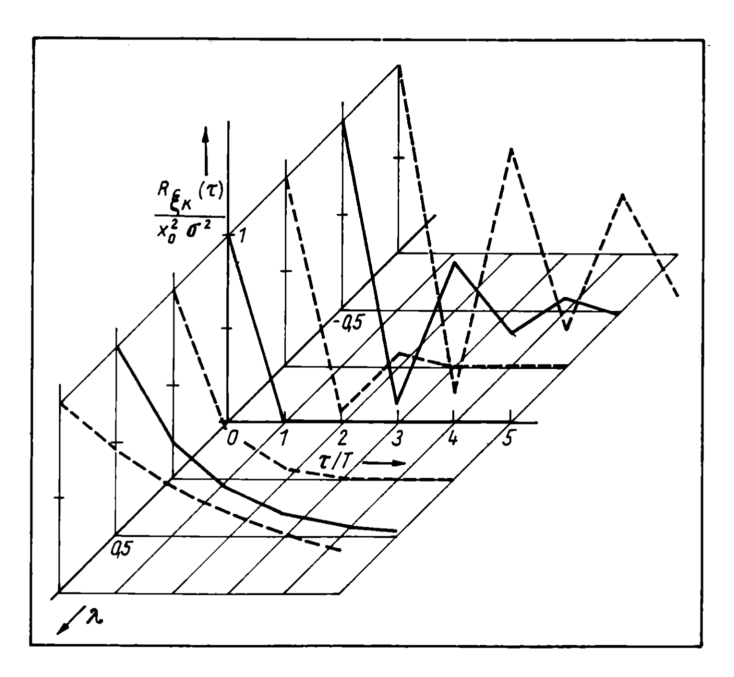


Bild 2. Autokorrelationsfunktion des binären Markovschen Impulsprozesses; $\lambda = 1 - (\alpha + \beta)$

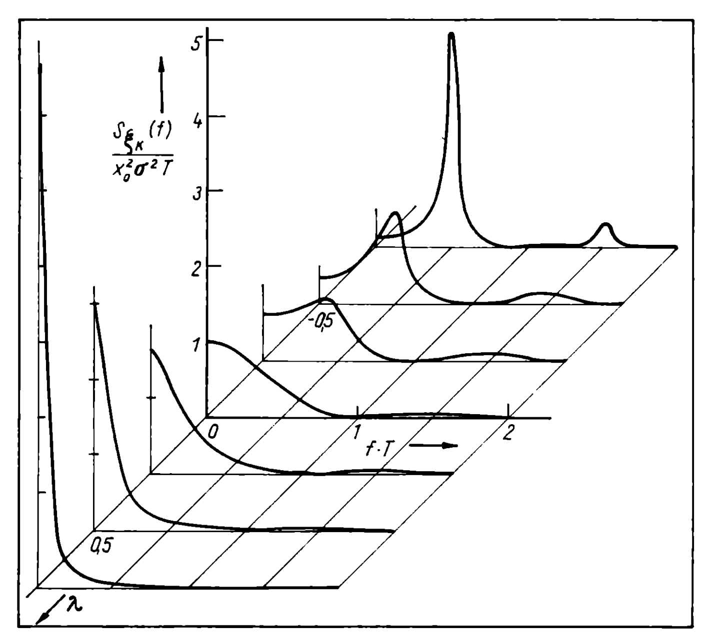


Bild 3. Leistungsdichtespektrum des binären Markovschen Impulsprozesses; $\lambda = 1 - (\alpha + \beta)$

$$\boldsymbol{P}_1 = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \quad (35)$$

gewählt, da dieser Spezialfall besondere Bedeutung in der digitalen Signalübertragung und -verarbeitung hat.

Zunächst werden die Eigenwerte berechnet. Sie ergeben sich gemäß (27) aus der Lösung der algebraischen Gleichung 2. Grades

$$\det\begin{pmatrix} 1-\alpha-\lambda & \beta \\ \alpha & 1-\beta-\lambda \end{pmatrix} = 0 \quad (36)$$

zu $\lambda_1 = 1$ und $\lambda_2 = 1 - (\alpha + \beta)$. Damit erhält man die potenzierte Eigenwertmatrix

$$\underline{\Lambda}^n = \operatorname{diag}(1, \lambda^n), \quad \lambda = 1 - (\alpha + \beta). \quad (37)$$

Mit der Matrix der Adjunkte

$$\mathbf{T} = \begin{pmatrix} -\beta & \alpha \\ -\alpha & -\alpha \end{pmatrix} \quad (38)$$

ergibt sich aus (24) die Übergangsmatrix nach n Schritten

$$\mathbf{P}_n = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta + \alpha\lambda^n & \alpha[1-\lambda^n] \\ \beta[1-\lambda^n] & \alpha + \beta\lambda^n \end{pmatrix}, \quad n \geqq 0. \quad (39)$$

Der Grenzübergang $n \to \infty$ führt zur Matrix

$$\mathbf{P} = \lim_{n\to\infty} \mathbf{P}_n = \begin{pmatrix} p_1 & p_2 \\ p_1 & p_2 \end{pmatrix} = \frac{1}{\alpha+\beta}\begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} \quad (40)$$

der ergodischen Grenzwahrscheinlichkeiten p_l. Mit den angeführten Gleichungen errechnet sich aus (33) die gesuchte Autokovarianzfolge

$$\mathbf{B}(n) = \sigma^2 \lambda^{|n|}, \quad \lambda = 1 - (\alpha + \beta), \quad (41)$$

worin σ^2 die Streuung

$$\sigma^2 = \frac{\alpha \cdot \beta}{\alpha + \beta}(\xi_1 - \xi_2)^2 \quad (42)$$

der Zufallsgröße ξ_l bezeichnet. Deren linearer Mittelwert folgt aus

$$m_1 = \frac{1}{\alpha+\beta}(\beta\,\xi_1 + \alpha\,\xi_2). \quad (43)$$

Mit der gefundenen Autokovarianzfolge erhält man schließlich aus den in Abschnitt 2.1 angegebenen Gleichungen für den binären Markovschen Impulsprozeß die Autokorrelationsfunktion

$$R_{\xi k}(\tau) = \frac{\sigma^2}{T} \sum_{n=-\infty}^{\infty} \lambda^{|n|} \cdot R_x(\tau - n\,T), \quad \lambda = 1 - (\alpha + \beta) \quad (44)$$

$$R_{\xi d}(\tau) = \frac{m_1^2}{T} \sum_{n=-\infty}^{\infty} R_x(\tau - n\,T) \quad (45)$$

und das Leistungsdichtespektrum

$$S_{\xi k}(f) = \frac{\sigma^2}{T}|X(f)|^2 \frac{1-\lambda^2}{1 - 2\lambda\cos(2\pi f T) + \lambda^2}, \quad \lambda = 1 - (\alpha + \beta) \quad (46)$$

$$S_{\xi d}(f) = \frac{m_1^2}{T^2} \sum_{n=-\infty}^{\infty} \left|X\left(\frac{n}{T}\right)\right|^2 \delta\left(f - \frac{n}{T}\right). \quad (47)$$

Man erhält (46), indem in (13) der Kosinus als Realteil einer e-Funktion mit imaginären Exponenten aufgefaßt wird und mit $\lambda < 1$ eine Potenzreihenentwicklung mit angebbarem Grenzwert möglich ist.

Für den rechteckförmigen Elementarimpuls der Amplitude x_0 und der Breite $t_i = T$ ist der Verlauf von Autokorrelationsfunktion und Leistungsdichtespektrum bei verschiedenen λ-Werten in den Bildern 2 und 3 dargestellt.

4. Zusammenfassung

Ausgehend von der Klasse periodischer stochastischer Impulsprozesse als hinreichendes Modell digitaler Signale wurde ein Überblick zur Berechnung von Autokorrelationsfunktion und Leistungsdichtespektrum digitaler Signale gegeben. Das verwendete mathematische Modell ermöglicht ein übersichtliches Berechnungsverfahren. So ist der Verlauf von Korrelationsfunktion und Spektralfunktion durch einen von der Impulsform bestimmten Formanteil und einen von der Impulsstruktur abhängigen Strukturanteil festgelegt. Das eigentliche Problem besteht in der Ermittlung der Autokovarianzfolge des Strukturanteils. Trivial ist diese Aufgabe für den Spezialfall statistisch voneinander unabhängiger Impulse. Eine allgemeine geschlossene Lösung gelingt für den Fall der Modellierung der statistischen Bindungen durch eine Markov-Kette. Die aufgeführte Lösung wurde am Beispiel des praktisch relevanten Binärsignals veranschaulicht.

NaA 16

Literatur

[1] *Lewin, B. R.:* Theoretische Grundlagen der statistischen Radiotechnik. Bd. 1. Moskau: Sowjetskoje radio 1974

[2] *Kaden, H.:* Theoretische Grundlagen der Datenübertragung. München, Wien: R. Oldenbourg 1968

[3] *Wehrmann, W.:* Einführung in die stochastisch-ergodische Impulstechnik. München, Wien: R. Oldenbourg 1973

[4] *Oehlen, H.; Brust, G.:* Spektrale Energie- und Leistungsdichte technisch interessanter Impulsformen und Impulsfolgen. AEÜ 22 (1968) 2, S. 79–86

[5] *Franks, L. E.:* Signal Theory. Englewood Cliffs. (N. Y.): Prentice-Hall Inc. 1969

[6] *Kreß, D.:* Theoretische Grundlagen der Übertragung digitaler Signale. Berlin: Akademie-Verlag 1979

[7] *Nuszkowski, H.:* Analyse von Markov-Ketten mit Signalflußgraphen. Wiss. Z. Techn. Univ. Dresden 24 (1975) 2, S. 411–414

[8] *Fisz, M.:* Wahrscheinlichkeitsrechnung und mathematische Statistik. Berlin: VEB Deutscher Verlag der Wissenschaften 1978

Dr. sc. techn. *Hans-Joachim Thomanek*, Ingenieurhochschule Mittweida, Sektion Informationselektronik, Platz der DSF 17, Mittweida, 9250

OPTIMOD – Ein universelles Programmsystem zur Unterstützung des Modellierungsprozesses im Schaltkreisentwurf

K. Freese; D. Warning, Berlin

Mitteilung aus dem Zentrum für Forschung und Technologie Nachrichtenelektronik Berlin

Die Netzwerkanalyse von analogen und digitalen Schaltungskomplexen hat sich zur Sicherung der Funktionsfähigkeit von integrierten Schaltkreisen vor der Präparation durchgesetzt. Dabei nimmt die Modellierung des elektrischen Verhaltens der aktiven integrierten Bauelemente eine Schlüsselstellung für die Güte des Analyseergebnisses ein. Dieser Modellierungsprozeß wird durch folgende Faktoren erschwert:

- Erhöhung der Komplexität der innerelektronischen Vorgänge durch die fortschreitende Miniaturisierung der Bauelementestrukturen
- Erweiterung der Analyse im Zeit- und Temperaturbereich
- bei der Wahl der Modellapproximation sind Rechenzeit- und Genauigkeitsanforderungen zu beachten.

Eine effektive Gestaltung dieses Prozesses ist nur bei Zugrundelegung einer durchgängigen Datenbasis von Technologiedaten über verdichtete elektrische Daten bis zur Modellbibliothek der Analyseprogramme möglich. Die Transformation der elektrischen Daten in die spezifischen Modelldaten (Parameter) hat entscheidenden Einfluß auf die für die Netzwerkanalyse relevante Modellgenauigkeit. Dabei ist es möglich, neben der eigentlichen Parameterextraktion auch Informationen über die Eigenschaften des Modellansatzes zu erhalten.

1. Einordnung des Programmsystems in den Prozeß des Schaltkreisentwurfs

Ziel des Programmsystems OPTIMOD ist es, den Entwicklern von Schaltkreisen aktuelle und zuverlässige Parametersätze für Modelle integrierter Bauelemente zur Verfügung zu stellen und die Entwicklung neuer Modellansätze zu unterstützen.
Die in den gebräuchlichen Analysemodellen vorhandene Redundanz zwischen verschiedenen Parametern, wie auch das nichtlineare Verhalten miniaturisierter Bauelemente führt beim Einsatz von globalen Optimierungsverfahren zum Auftreten von Nebenminima [1] [14] [15]. Dieses Problem ist durch eine stufenweise Bearbeitung teilweise lösbar. Adaptive Suchverfahren haben dabei die Aufgabe, aus einem verhältnismäßig großen Parameterbereich unter geringen Anforderungen an das Fehlerkriterium Startwerte für die genauer arbeitenden Gradientenverfahren zur Verfügung zu stellen. Bild 1 zeigt die Struktur des Programmsystems, die durchgängige Datenbasis sowie die Ankopplung an den Schaltkreisentwurf.

2. Mathematisches Verfahren

Bei der Wahl des Optimierungsverfahrens ist zu berücksichtigen, daß eine analytische Berechnung der 1. und 2. Ableitungen vermieden werden sollte, um eine universelle und einfache Anpassung des Optimierungsprogramms an die jeweilige Aufgabenstellung zu erreichen. Trotzdem sollte eine große Genauigkeit und eine möglichst große Robustheit gegenüber schlecht konditionierten Problemstellungen bzw. schlechten Startwerten erreicht werden. Als Optimierungsfunktion wird das mittlere Fehlerquadrat über alle gegebenen Meßpunkte verwendet

$$F(\underline{x}) = \sum_{i=1}^{m} [y_i(\underline{x}) - y_i^*]^2 = \sum_{i=1}^{m} [f_i(\underline{x})]^2 = \|\underline{f}(\underline{x})\|_2^2 = \underline{f}^T \cdot \underline{f} \qquad (1)$$

y_i^* Meßwert im i-ten Meßpunkt
$y_i(\underline{x})$ Modellwert mit Parametervektor $\underline{x}$ im i-ten Meßpunkt
$f_i(\underline{x})$ Fehler im i-ten Meßpunkt
$\underline{f}(\underline{x})$ Fehlervektor.

Zwei Varianten der Wichtung sind möglich. Zum einen kann eine Größe $y_{\min}$ angegeben werden. Für Meßwerte $y_i^* \geqq y_{\min}$ wird mit dem relativen Fehler gerechnet. Bei Meßwerten $y_i^* < y_{\min}$ geht dagegen der absolute Fehler, skaliert mit $y_{\min}$, in die Rechnung ein. Fehler in diesen Meßwerten haben damit weniger Einfluß auf die Optimierungsfunktion. So kann eine untere Grenze in der Genauigkeit der Meßwerte oder im Gültigkeitsbereich des zu optimierenden Modells berücksichtigt werden.
Weiterhin ist es möglich, für jeden einzelnen Meßpunkt i eine Wichte w anzugeben. Diese Art der Wichtung ist zum Beispiel notwendig für Abschnitt der Meßfunktion, deren Kurvenverläufe genauer approximiert werden sollen.
Aus (1) wird dann

$$F(\underline{x}) = \sum_{i=1}^{m} \left[\frac{y_i(\underline{x}) - y_i^*}{\max(y_i^*, y_{\min})} \cdot w_i\right]^2. \qquad (2)$$

Ziel der Optimierung ist es nun, den Parametervektor $\underline{x}$ so zu bestimmen, daß $F(\underline{x})$ minimal wird.
Aufgrund der in der Einleitung genannten Problemstellung wurde als Optimierungsverfahren der *Levenberg-Marquardt*-(L–M-)Algorithmus verwendet, der Vorteile der Methode des steilsten Abstiegs und des *Gauß-Newton*-Verfahrens vereinigt. Der L–M-Algorithmus verbessert den Parametervektor $\underline{x}$ nach folgender Iterationsvorschrift [1]:

$$\underline{x}_{k+1} = \underline{x}_k - [\mathrm{J}^T \mathrm{J} + \lambda_k \cdot \mathrm{diag}(\mathrm{J}^T \mathrm{J})]^{-1} \cdot \mathrm{J}^T \underline{f}(\underline{x}_k) \qquad (3)$$

λ_k Marquardt-Parameter
J Jacobi-Matrix
diag $(\mathrm{J}^T \mathrm{J})$ Diagonal-Matrix mit den Diagonalen-Elementen von $(\mathrm{J}^T \mathrm{J})$.

Die Jacobi-Matrix des Fehlervektors ist die $(i \cdot j)$-Matrix, deren Elemente jeweils die 1. Ableitung nach Parameter j im Meßpunkt i sind. Zur numerischen Berechnung dieser Ableitungen werden im Programm zwei Varianten angeboten. Im allgemeinen genauer, aber aufwendiger ist die Näherung über zentrale Differenzen

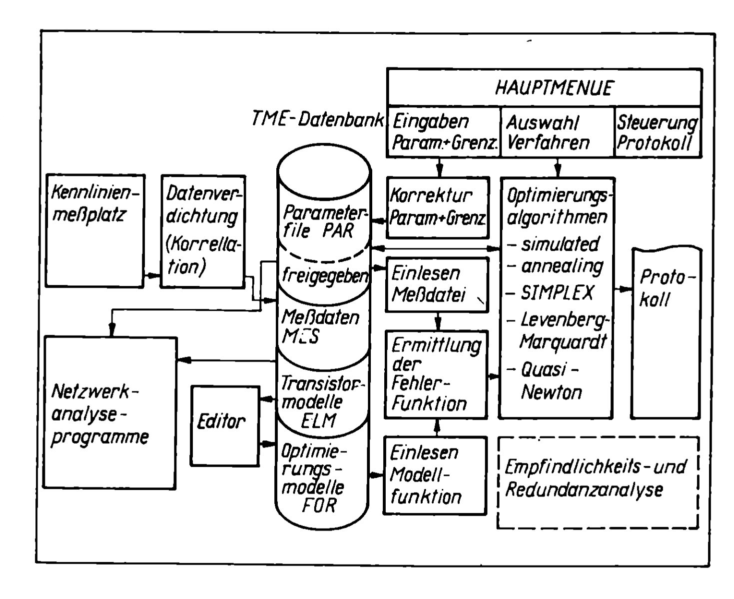
$$\mathrm{J}_{ij}(\underline{x}_k) \approx \frac{y_i(\underline{x}_k + h_j^k \cdot \underline{e}_j) - y_i(\underline{x}_k - h_j^k \cdot \underline{e}_j)}{2 \cdot h_j^k}. \qquad (4a)$$

In einzelnen Fällen ist die Näherung mit Vorwärtsdifferenzen günstiger

$$\mathrm{J}_{ij}(\underline{x}_k) \approx \frac{y_i(\underline{x}_k + h_j^k \cdot \underline{e}_j) - y_i(\underline{x}_k)}{h_j^k} \qquad (4b)$$

e_j j-ter Spaltenvektor der Einheitsmatrix
h_j^k kleiner Schritt in Richtung des Parameters j.

Bild 1. Programmsystem zur Parameterextraktion



	Ausgangswerte	opt. Werte
K20	0.75	0.69
K21	0.23	0.2
VBK	0.83	− 0.71
K3	0.189	0.22
K4	0.0424	0.0431
K5	0.86	0.74
BETA	0.0252	0.0319
UTO	0.8	0.76
2VF	0.65	0.64
DB	1.6	2.13
DL	− 2.0	− 1.45
LTYP	3.0	2.3
X	0.4	− 0.194
Y	0.8	0.621
Z	0.8	1.04
DK3	0.261	0.213
DK5	0.14	− 0.067
DBET	0.0119	0.0131
DVBK	1.1	1.97

Tafel 1. Parameter Enhancementtransistor

Die angemessene Wahl von h_j^k steht in engem Zusammenhang mit der gegebenen Problemstellung und der vorhandenen Maschinengenauigkeit. Sie hat wesentlichen Einfluß auf die Qualität der Näherung der *Jacobi*-Matrix und damit auch auf die schnelle Konvergenz des Verfahrens.
Hier wird h_j^k berechnet nach [2]:

$$h_j^k = \min\left[\max \|\underline{y}_i(\underline{x}_k)\|, \delta_j^k\right]$$

mit

$$\delta_j^k = \begin{cases} 10^{-10} & \text{für } |x_j^k| < 10^{-6} \\ 0.0001 \cdot |x_j^k| & \text{für } |x_j^k| \geqq 10^{-6}. \end{cases} \tag{5}$$

Eine weitere Verringerung des Rechenaufwands wird im Programm dadurch erreicht, daß die *Jacobi*-Matrix erst nach einer Anzahl von Iterationsschritten neu berechnet wird. Bis dahin wird nach jedem Iterationsschritt mit einer Aufdatierungsformel [1]:

$$\mathrm{J}_{k+1} \approx \mathrm{J}_k + \frac{(\Delta f_k - \mathrm{J}_k \cdot \Delta \underline{x}_k) \cdot \Delta \underline{x}_k^T}{\|\Delta \underline{x}_k\|_2^2} \tag{6}$$

berechnet.
Zu Beginn der Optimierung wird bei schlechtem Startwert für $\underline{x}$ ein großes λ gewählt — man erreicht wie bei der Methode des steilsten Abstiegs große Verbesserungen der Optimierungsfunktion. Bei den weiteren Iterationsschritten wird λ verkleinert, so daß in der Nähe des Minimums der Algorithmus zum *Gauß-Newton*-Verfahren tendiert und dessen rasche Konvergenz erreicht.
Die Berechnung von λ für einen Iterationsschritt erfolgt nach [3]:

$$\lambda = \frac{2 \cdot \underline{g}^T \cdot \underline{v}}{F(\underline{x})} \quad \text{mit} \quad \begin{aligned} \underline{g} &= \mathrm{J}^T \cdot \underline{f}(\underline{x}) \\ \underline{v} &= \left[g_1 \cdot (\mathrm{J}^T \cdot \mathrm{J})_{11}^{-1}, \ldots, g_j \cdot (\mathrm{J}^T \cdot \mathrm{J})_{jj}^{-1}\right]. \end{aligned} \tag{7}$$

Verringert die Iteration mit diesem λ den Wert von $F(\underline{x})$ nicht, dann wird der Iterationsschritt mit einer jeweiligen Vergrößerung von λ so lange wiederholt, bis eine Reduzierung erreicht ist. Ein weiteres wichtiges Problem ist die Berücksichtigung linearer Restriktionen für die zu optimierenden Parameter. Dadurch kann erreicht werden, daß Parameter nur innerhalb von entsprechend der Aufgabenstellung sinnvollen Intervallen optimiert werden.
Überschreiten in einem Iterationsschritt ein oder mehrere Parameter diese Grenzen, so wird zunächst versucht durch Vergrößerung von λ einen Parametervektor innerhalb der Grenzen zu finden, der $F(\underline{x})$ reduziert. Dann werden die Parameter festgesetzt, deren Grenzen verletzt wurden, d.h., sie werden im nächsten Iterationsschritt nicht in die Optimierung einbezogen. Dazu wird eine $(n - q \cdot n)$-Projektionsmatrix D (n Parametergesamtzahl, q Anzahl der festgesetzten Parameter) gebildet, indem aus der $(n \cdot n)$-Einheitsmatrix für jeden Parameter j der festgesetzt werden soll der j-te Zeilenvektor herausgenommen wird.
Die Iterationsvorschrift (3) erhält damit folgende Form:

$$\underline{x}_{k+1} = \underline{x}_k - \mathrm{D}^T \cdot \{\mathrm{D} \cdot [\mathrm{J}^T \mathrm{J} + \lambda_k \cdot \mathrm{diag}\,(\mathrm{J}^T \mathrm{J})] \cdot D^T\}^{-1} \cdot \mathrm{D} \cdot \mathrm{J}^T \cdot f(\underline{x}_k). \tag{8}$$

Festgesetzte Parameter werden bei folgenden Iterationsschritten nacheinander wieder frei, um zu überprüfen, ob sie bei inzwischen verändertem Parametervektor $\underline{x}$ immer noch in Richtung der Grenzen tendieren oder wieder voll in die Optimierung einbezogen werden können.
Der dritte Komplex, neben der Reduzierung von $F(\underline{x})$ und der Beachtung der Restriktionen, der nach jedem Iterationsschritt getestet werden muß, ist die Erfüllung von Abbruchbedingungen. Mit diesen Bedingungen wird festgelegt, wann der aktuelle Parametervektor mit ausreichender Genauigkeit ein Minimum für $F(\underline{x})$ erreicht hat und mit weiteren Iterationen nur noch unwesentlich verbessert werden könnte.

3. Programmimplementierung

Das vorgestellte Verfahren wurde in PASCAL codiert. Es besteht aus mehreren Objekt-Modulen, die mit dem Objekt-Modul des zu optimierenden Modells verbunden werden. Die Modelle können beliebige Funktionen sein. Die Anpassung der Funktion an die Schnittstelle des Optimierungsprogramms wurde auf einfache Parameterübergaben und -zuweisungen beschränkt, so daß im Originalmodell nur wenige Änderungen vorzunehmen sind. Das gesamte Programm läuft dialogorientiert ab. Der Hauptteil der Eingaben besteht in der Zuweisung der Startwerte und ggf. zugehöriger Grenzen der Parameter. Dem Nutzer ist es möglich, über Eingaben die Optimierung seinen Wünschen entsprechend zu modifizieren, wobei jeweils eine Standardvariante angeboten wird. Das betrifft

- $y_{\min}$
- Wichten für die einzelnen Meßpunkte
- Anzahl der Iterationsschritte mit Aufdatierung der *Jacobi*-Matrix bis zur Neuberechnung der *Jacobi*-Matrix
- Rechnung mit zentralen oder Vorwärtsdifferenzen bei der Näherung der 1. Ableitungen.

Ebenfalls variiert werden kann der Umfang der Auswertung, von der Protokollierung des optimierten Parametervektors bis zur vollständigen Gegenüberstellung von Meßwerten, originalen und optimierten Modellwerten.

4. Praktische Erfahrungen

Das Programmsystem OPTIMOD wurde bei der Modellierung des statischen elektrischen Verhaltens von MOS-Transistoren (nSGT) eingesetzt. Die folgenden Beispiele zeigen wie breit das Programm eingesetzt werden kann.

4.1. *Stromprozeduren für Enhancementtransistoren*

Für das zur Simulation von digitalen Schaltungen genutzte Modell nach [9] [10] wurde an 5 Transistoren (B/L: 30/4,5, 30/5, 30/6, 30/30, 6/6) mit einem Datensatz von ≈ 500 Meßpunkten ($UDS = 0$ bis 6 V, $UGS = 0$ bis 5 V, $UBS = -0{,}25$ bis 4,5 V) ein kompletter Parametersatz ermittelt. Die Ergebnisse sind in Tafel 1 und im Bild 2 dargestellt.
Die gemittelte Anpaßgenauigkeit beträgt 7,4 % und wird in diesem einfachen Modell durch den schwellspannungsnahen Arbeitsbereich sowie den Hochstrom-Bereich bestimmt.
Aufgrund der vorhandenen Redundanz zwischen einigen Parametern (z.B. K20, K21, VBK) wurde in 3 Rechengängen optimiert. Der gesamte CPU-Zeitbedarf lag bei ≈ 21 min.
Für das LEVEL2-Modell nach [16] konnten an 750 Meßpunkten der nSGT 3-Technologie folgende Parameter bestimmt werden:

VTO	: 0,63 V	XJ	: 4.5E-7 m
KP	: 3,20E-5 A/V²	LD	: 8.0E-7 m
GAMMA	: 0.57 $\sqrt{\mathrm{V}}$	*UO*	: 695 cm/Vs
PHI	: 0.73 V	*UCRIT*	: 6.7E+4 V/cm
TOX	: 5.9E-8 m	*UEXP*	: 0.12
NSUB	: 1.5E+15 cm⁻³	*VMAX*	: 4.9E+4 m/s
NFS	: 1.2E+11 cm⁻²	*NEFF*	: 3.5

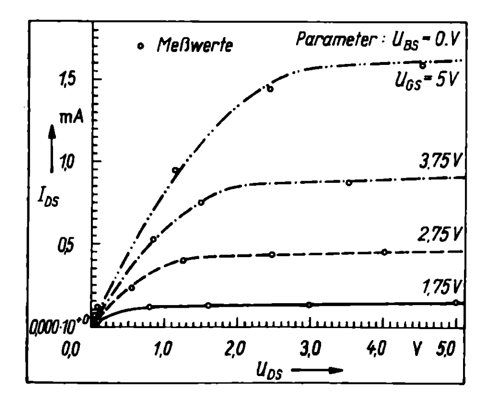


Bild 2. Ausgangskennlinienfeld (B/L = 30/5)

Die vorgestellte Verfahrensweise kann aufgrund ihrer Einfachheit zur Parameterextraktion an technologietypischen Transistoren genutzt werden und bildet somit eine Möglichkeit zur Technologiekontrolle. — Es ist aber auch möglich Transistorparameter an speziellen ausgewählten Scheiben zu ermitteln, und mit Hilfe der Simulation den Prozeß der Fehlersuche zu unterstützen.

4.2. Schwellspannungsmodelle für Enhancementtransistoren

Zur Modellierung der Geometrie- und Spannungsabhängigkeit der Schwellspannung von Enhancementtransistoren wurden 2 Modelle an einem Datensatz $UT = f(B, L, USB, UDS)$ mit etwa 60 Meßpunkten untersucht.

Für folgenden empirischen Modellansatz nach [11] wurden mit dem beschriebenen Verfahren Parameter extrahiert:

$$UT = UFB + 2 * Vf + K2 * \sqrt{USB + 2 * Vf} + \left(K2I + \frac{K2B}{B + DB} - \frac{K2L}{L + DL}\right) * (USB + 2 * Vf) + K2D * VD \quad (9)$$

mit $UFB = 0.9$, $Vf = 0.33$, $K2 = 1.3$, $K2I = -0.21$, $K2B = 0.28$, $K2L = 0.16$, $K2D = 0.1$

Die protokollierte mittlere Genauigkeit ist 2,9%.

Für das geometrisch-technologische Modell nach [12], das auf der Beschreibung der vom Gate beeinflußten Ladungsmenge beruht, wurden 9 technologische und geometrische Parameter innerhalb ihrer Schwankungsbereiche, sowie 2 empirische Parameter bestimmt. Die resultierende mittlere Genauigkeit ist 7,5% und wird durch die im Modell nicht berücksichtigte inhomogene Kanaldotierung begrenzt. Diese Inhomogenität zeigt sich im Kennlinienfeld $UT = f(USB)$ Bild 3 mit einem Abknicken bei $USB = 0.9$ V.

4.3. Substratstrommodellierung — Multiplikationsfaktor

Substratströme treten insbesondere bei n-Kanal Transistoren im Sättigungsbereich auf. Diese durch Stoßionisation verursachte Stromkomponente führt zu einer Erhöhung des Drainstroms und kann für die Berechnung analoger Schaltungen bzw. bei der Analyse von Latch-up-Vorgängen in der CMOS-Technik von Bedeutung sein. Für folgenden Modellansatz [13] wurden die Parameter aus einem Datensatz von etwa 400 Meßpunkten bestimmt:

$$ISUB = IDS \cdot \frac{2 \cdot EX^{A} \cdot AI}{K4SU \cdot BI} \cdot \exp\left(-\frac{BI}{EX}\right) \quad (10)$$

mit $EX = F * \sqrt{K4SU(UDS - VSAT) + \left(\frac{EG}{F}\right)^2}$

$A = K1S * UDS^2 + K2S$

$AI = 373\,\mu m^{-1}$, $BI = 172\,\frac{V}{\mu m}$, $K4SU = 2.4$, $EG = 4 \cdot \frac{V}{\mu m}$

$F = 3.37$, $K1S = -0.01\,V^{-2}$, $K2S = 2.8$.

Das Ergebnis ist im Bild 4 dargestellt. Der nach der Parameterextraktion verbleibende Restfehler ist mit dem Auftreten von weiteren Durchbruchserscheinungen zu erklären, welche im Ansatz (10) nicht berücksichtigt sind.

5. Zusammenfassung

Das dargestellte Verfahren der Parameterextraktion ist ein wichtiges Glied im Prozeß der technologiegestützten Modellierung von integrierten Bauelementen für die Schaltkreissimulation. Aufgrund der einfachen Handhabbarkeit und der gewählten Optimierungsmethoden ist ein Einsatz von OPTIMOD für technologische Problemstellungen bzw. für den Schaltungsentwurf (z.B. Filter) denkbar.

Mit einer Erweiterung um eine anschließende Empfindlichkeits- und Redundanzanalyse wird das Programmsystem OPTIMOD zu einem leistungsfähigen Software-Werkzeug zur Modellentwicklung im Schaltkreisentwurf.

NaA 28

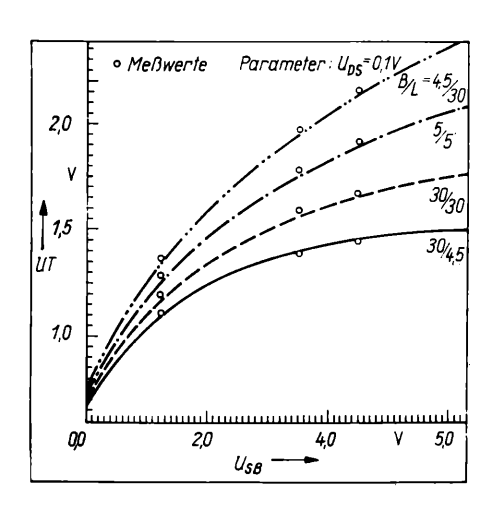


Bild 3. Schwellspannungsmodelle

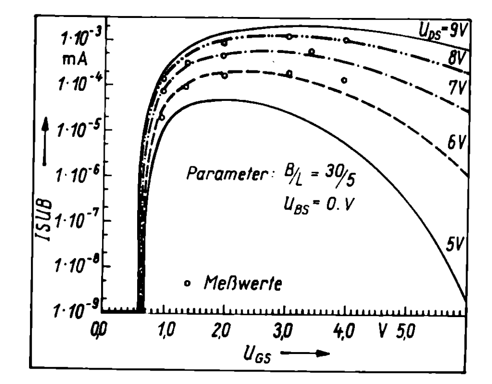


Bild 4. Substratstromverhalten Enhancementtransistor

Literatur

[1] *Doganis, K.; Scharfetter, D. L.:* General Optimization and Extraction of IC Device Model Parameters. IEEE Trans. on Electron Devices, 30 (1983) 9, S. 1219—1228

[2] *Brown, K. M.; Dennis, J. E.:* Derivate Free Analogues of the Levenberg-Marquardt and Gauss Algorithms for Nonlinear Least Squares Approximation. In: Numerische Mathematik. Band 18, 4. Berlin: Springer-Verlag 1972

[3] *Davies, M.; Whitting, I. J.:* A Modified Form of Levenberg's Correction. In: *Lootsma, F. A.:* Numerical Methods for Nonlinear Optimization. 2. Auflage. London: Academic Press 1976

[4] *Krabs, W.:* Einführung in die lineare und nichtlineare Optimierung für Ingenieure. 1. Auflage, Leipzig: Teubner Verlagsgesellschaft 1983

[5] *Psenicnyi, B. N.; Danilin, J. M.:* Numerische Methoden für Extremalaufgaben. Berlin: Deutscher Verlag der Wissenschaften 1982

[6] *Entenmann, W.:* Optimierungsverfahren. UTB 539. Heidelberg: Hüthig-Verlag 1976

[7] *More, J. J.:* The Levenberg-Marquardt Algorithm: Implementation and Theorie. *Watson, G. A.:* Numerical Analysis. Lecture Notes in Mathematics 630. Berlin: Springer-Verlag 1978

[8] *Zurmühl, R.; Falk, S.:* Matrizen und ihre Anwendungen. Teil 1, 5. Auflage. Berlin: Springer-Verlag 1984

[9] *Krauß, M.:* Ein entwurfsorientiertes, rechnergerechtes Schaltungsmodell für den MOS-Anreicherungstransistor. Nachrichtentech., Elektron. 28 (1978) 1, S. 10—11

[10] *Posdziech, G.:* Empirisches Modell des Enhancement-Transistors mit kurzem Kanal. Nachrichtentech., Elektron. 30 (1980) 3, S. 111—112

[11] *Silburt, A. L.; Foss, R. C.; Petrie, W. F.:* An Efficient MOS Transistor Model For Computer Aided Design. IEEE Trans. on Computer Aided Design 3 (1984) 1, S. 104—110

[12] *Jäntsch, O.:* A Geometrical Model of the Threshold. Voltage of Short and Narrow Channel MOSFET's. Solid State Electronics, 25 (1982) 1, S. 59—61

[13] *Schumacher, K.:* Modellierung, Entwurf und Untersuchung von MOS-Operationsverstärkern für die Großintegration. Univ. Dortmund 1978, Diss. A

[14] *Brandt, K.; Madas, I.; Mirbeth, M.; Zapf, H. L.:* Parameterbestimmung für MOS-Modelle und Rechnergestützte Dokumentation von Parametersätzen für Simulationszwecke. NTG-Fachberichte, Bd. 82, (1983) S. 70—73

[15] *Conway, P.; Cahill, C.; Lane, W. A.; Lidholm, S. U.:* Extraction of MOSFET Parameters Using the Simplex Direct Search Optimization Method. IEEE Trans. on CAD, 4 (1985) 4, S. 694—698

[16] *Vladimirescu, A.; Liu, S.:* The Simulation of MOS Integrated Circuits using SPICE2. Memo. No. UCB/ERL M80/7, University of California, Berkeley, 1980

Dipl.-Ing. *Dietmar Warning*, VEB Elektroprojekt Anlagenbau Berlin, Stammbetrieb des Kombinat Automatisierungsanlagenbau, Zentrum für Forschung und Technologie Abtlg. CES, Rhinstr. 100, Berlin, 1140

Dipl.-Ing. *Kai Freese*, Technische Universität Karl-Marx-Stadt Sektion Informationstechnik, PSF 964, Karl-Marx-Stadt, 9010

Großsignalverhalten der Avalanchefotodiode beim Einsatz in Breitbandempfängern für Analogsignale

D. Rau, KDT, Berlin

Mitteilung aus der Humboldt-Universität zu Berlin, Sektion Elektronik

Obwohl ursprünglich für den Einsatz in digitalen LLNÜ[1])-Systemen vorgesehen, findet die Avalanchefotodiode APD auch bei optischen Empfängern für Analogsignale Verwendung. Dabei müssen jedoch folgende wesentliche Einschränkungen beachtet werden:

1. Wie bereits u.a. in [1] beschrieben, wird durch den stochastischen Vervielfachungsprozeß im Avalanchegebiet der Diode ein zusätzliches weißes Rauschen erzeugt, das gegenüber M überproportional anwächst und durch $F(M)$ hinreichend genau beschrieben werden kann [2] [3]. Aufgrund dieser Eigenschaft sind die bei Multiplikationsfaktoren $M \gtrless 20$ erreichbaren Signal-Rausch-Abstände für breitbandige Analogsignale im allgemeinen zu gering, so daß M selten größer als etwa 10 gewählt werden kann.

2. Bedingt durch ihren Aufbau haben heute übliche Avalanchefotodioden bei kleinen Multiplikationsfaktoren (und damit kleiner Sperrspannung) eine z.T. erheblich höhere Sperrschichtkapazität als pin-Dioden, so daß sich sowohl bei Hoch- als auch Transimpedanzempfängern mit APD bei Multiplikationsfaktoren $\geqq 1$ eine deutliche Verringerung der oberen Grenzfrequenz nachweisen läßt.

Aufgrund der im Teilnehmer-Anschlußbereich verstärkt auftretenden variablen Feldlängen zwischen Sender und Empfänger müssen, falls zusätzlicher (u. U. hoher) Aufwand zur Verstärkungsregelung im optischen Empfänger bzw. zur Sendeleistungsreduzierung im Sender nicht möglich oder ökonomisch vertretbar ist, die eingesetzten Empfängerschaltungen großsignalfest sein, d.h. auch bei hohen Eingangslichtleistungen noch im linearen Betrieb arbeiten. Anderenfalls kommt es bei analogen Frequenzmultiplexsignalen zu Nebensprechen und Klirrverzerrungen. Speziell beim Einsatz von APD können auch bei Vervielfachungsfaktoren $M < 10$ und Eingangssignalen hoher Intensität oder mit hohem Gleichlichtanteil Verzerrungen auftreten, die durch eine dynamische Änderung des Multiplikationsfaktors der APD hervorgerufen werden.

1. Statische und dynamische Multiplikationsfaktorverschiebung

Bild 1 zeigt die für analoge Breitbandempfänger typische kapazitive Ankopplung des Verstärkers (Hoch- und Transimpedanzverstärker) an die APD. Der zum Erzeugen der Vorspannung notwendige Gleichstrompfad wird mit R_V realisiert. Es sind 2 Multiplikationsfaktorverschiebungen zu unterscheiden:

1.1. Die statische Multiplikationsfaktorverschiebung $\varkappa_s$

Sie wird durch die mittlere Eingangslichtleistung P_0 hervorgerufen. Der durch R_V und die APD fließende Strom

$$I_s = S\,P_0 M \tag{1}$$

bewirkt einen Spannungsabfall über R_V, so daß die an der APD wirklich anliegende Vorspannung

$$U = U_0 - I_s R_V \tag{2}$$

beträgt. Damit ist der wirkliche Vervielfachungsfaktor $M = f(U_0 - I_s R_V)$, es gilt also $M < M_0$, wenn $M_0 = f(U_0)$ als berechneter Vervielfachungsfaktor angenommen wird.
Aus

$$\frac{M_0}{M} = 1 + \varkappa_s \tag{3}$$

wird die statische Multiplikationsfaktorverschiebung

$$\text{mit} \quad \varkappa_s = \frac{M_0}{M} - 1 \tag{4}$$

definiert.

[1]) LLNÜ Lichtwellenleiter-Nachrichtenübertragung

1.2. Die dynamische Multiplikationsfaktorverschiebung $\varkappa_d(t)$

Sie wird durch die Empfangswechsellichtleistung $p_{in}(t)$ (Spitzenwert $\hat{P}$) verursacht. Das Wirkungsprinzip ist das gleiche wie unter 1.1. beschrieben; der wirksame Widerstand ergibt sich jedoch zu

$$R_e = R_V \,||\, R_{in} \tag{5}$$

und damit wird die über der APD liegende Spannung

$$u(t) = U - i_s(t)\, R_e; \tag{6}$$

$\varkappa_d(t)$ läßt sich somit definieren zu

$$\varkappa_d(t) = \frac{M}{M(t)} - 1 \tag{7}$$

wobei $M(t) = f(U - i_s(t)\, R_e)$ gilt.
$\varkappa_s$ und $\varkappa_d(t)$ lassen sich prinzipiell nach dem gleichen, im folgenden beschriebenen Verfahren berechnen. Für die Praxis interessant erscheinen dabei die Abhängigkeiten $\varkappa_s(P_0)$ bzw. $\varkappa_d(p_{in})$. Erstere kann zur Bestimmung von Korrekturfaktoren für die einzustellende Vorspannung der APD bei bekannter mittlerer Eingangslichtleistung verwendet werden, letztere u.a. zur Ermittlung von durch $\varkappa_d$ hervorgerufene Klirrfaktoren des Eingangssignals bei bekannter Signalstruktur. Für die folgenden Ausführungen seien statische und dynamische Multiplikationsfaktorverschiebungen zusammengefaßt mit $\varkappa$ sowie der wirksame Widerstand allgemein mit R bezeichnet.
Allgemein gilt entsprechend (1)

$$I_s = S P_{in} M \tag{8}$$

und damit

$$U = U_0 - I_s R = U_0 - \frac{S P_{in} M_0}{1 + \varkappa} R. \tag{9}$$

Es ergibt sich somit zunächst $P_{in}(\varkappa)$ mit

$$P_{in} = \frac{(U_0 - U)(1 + \varkappa)}{S M_0 R}. \tag{10}$$

Da U_0 eine Funktion von M_0 ist und U von M und damit auch von $\varkappa$ abhängt, muß an dieser Stelle eine Funktion $U(M)$ eingeführt werden. Für $M \gtrless 3$ gilt nach [4] mit hinreichender Genauigkeit

$$M = \frac{1}{1 - \left(\frac{U}{U_{BR}}\right)^n}. \tag{11}$$

Die experimentellen Konstanten U_{BR} (Durchbruchspannung der APD) und n sind von der verwendeten APD abhängig. Die in diesem Beitrag benutzten Werte der SP 104 wurden durch Approximation einer Meßkurve ermittelt. Danach sind $U_{BR} = 398{,}1$ V und $n = 6{,}982$.

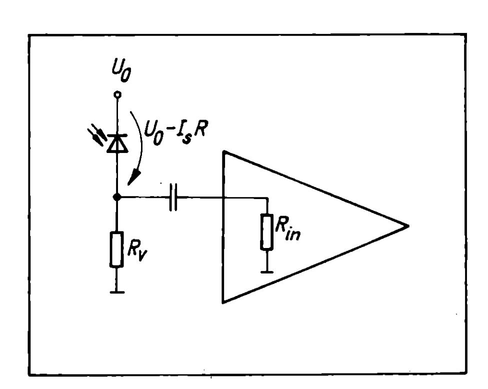


Bild 1. Kapazitive Ankopplung der APD bei Breitbandempfängern für Analogsignale

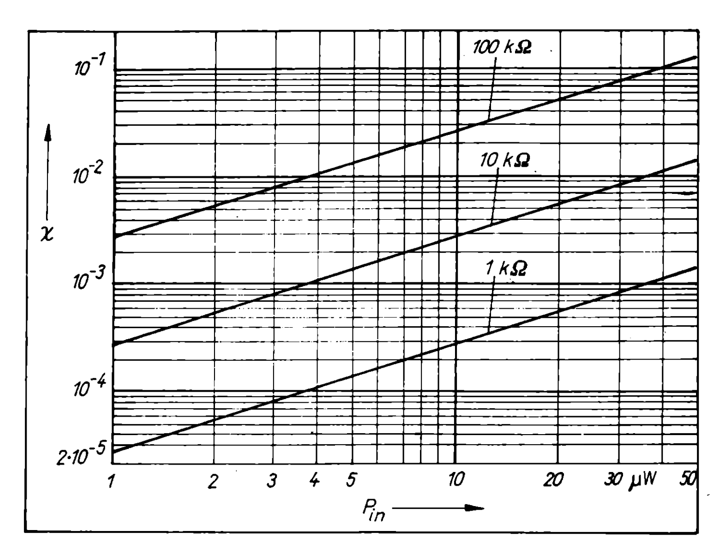


Bild 2. Multiplikationsfaktorverschiebung $\varkappa$ in Abhängigkeit von der Eingangslichtleistung P_{in}; Parameter R; $M_0 = 9$

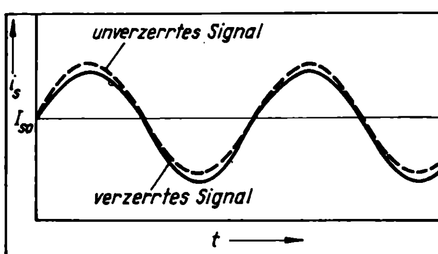


Bild 3. Nichtlineare Verzerrungen durch die dynamische Multiplikationsfaktorverschiebung $\varkappa_{\text{d}}$

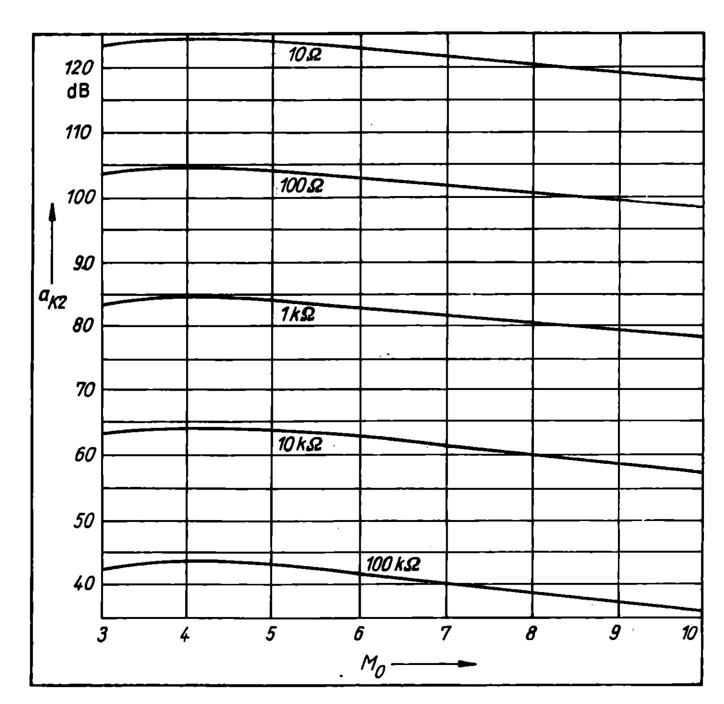


Bild 4. Klirrdämpfung $a_{\text{k}\,2}$ in Abhängigkeit vom Multiplikationsfaktor M_0; Parameter R_{e}

Durch Einsetzen von (11) in (10) wird

$$P_{\text{in}} S M_0 R = (1 + \varkappa) \left[U_0 - U_{\text{BR}} \left(1 - \frac{1}{M_0}\right)^n \left(1 - \frac{\varkappa}{M_0 - 1}\right)^n \right]. \quad (12)$$

Für $\varkappa \ll 1$ kann $\left(1 - \frac{\varkappa}{M_0 - 1}\right)^n \approx 1 - n \frac{\varkappa}{M_0 - 1}$ in erster Näherung gesetzt werden, so daß sich aus (12)

$$\varkappa^2 + \varkappa - \frac{P_{\text{in}} S M_0 R (M_0 - 1)}{n U_0} = 0 \quad (13)$$

ergibt. Hierbei interessiert nur die Lösung der quadratischen Gleichung, deren Absolutwert kleiner als 1 ist:

$$\varkappa = -\frac{1}{2} + \sqrt{\frac{n U_0 + 4 P_{\text{in}} S M_0 R (M_0 - 1)}{4 n U_0}}. \quad (14)$$

Damit läßt sich $\varkappa (P_{\text{in}})$ für $P_{\text{in}} = 1 \cdots 50$ µW mit $< 1\%$ Genauigkeit bestimmen, wie auf iterativem Wege berechnete Vergleichswerte zeigten. Bild 2 stellt die Abhängigkeit $\varkappa (P_{\text{in}})$ für verschiedene Werte des wirksamen Widerstandes R dar. Zu beachten ist, daß bei noch größeren Widerstandswerten als den hier angegebenen bzw. für große Eingangslichtleistungen (> 100 µW) oder großem Multiplikationsfaktor (> 20) die Multiplikationsfaktorverschiebung nicht mehr sehr viel kleiner als 1 ist und (14) damit falsche Ergebnisse liefert.

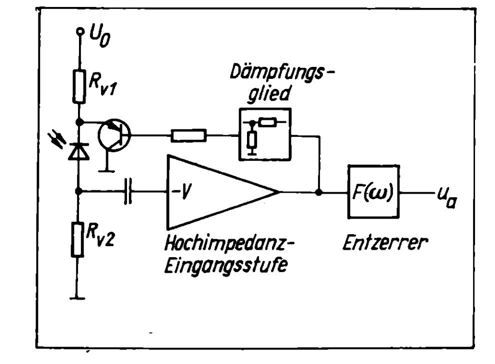


Bild 5. Prinzipschaltung zur dynamischen Stabilisierung der APD-Vorspannung bei Breitbandempfängern für Analogsignale

2. Klirrdämpfungsbestimmung

Wenn die Größe der dynamischen Multiplikationsfaktorverschiebung $\varkappa_{\text{d}}$ bekannt ist, läßt sich ermitteln, ob durch den Einsatz der APD nichtlineare Verzerrungen des Empfangssignals entstehen und wie hoch diese sind.

Bild 3 veranschaulicht am einfachsten Fall eines Eingangssignals (eine unmodulierte Trägerfrequenz) die durch die dynamische Multiplikationsfaktorverschiebung möglichen nichtlinearen Verzerrungen. Die obere Halbwelle des Signals wird gestaucht, während die untere Halbwelle gestreckt wird. Dadurch kommt es zur Ausbildung von Oberwellen, die bei entsprechender Größe evtl. vorhandene Nachbarkanäle im Frequenzmultiplexsignal stören.

Zur Berechnung der Klirrdämpfung $a_{\text{k}\,2}$ wird von (8) ausgegangen:

$$i_{\text{s}} = \frac{p_{\text{in}} S M_0}{1 + \varkappa_{\text{d}}}. \quad (15)$$

Mit (14) ergibt sich

$$i_{\text{s}} \approx \frac{2 p_{\text{in}} S M_0 \sqrt{n U_0}}{\sqrt{n U_0} + \sqrt{n U_0 + 4 p_{\text{in}} S M_0 R_{\text{e}} (M_0 - 1)}}. \quad (16)$$

Die Funktion $i_{\text{s}} (p_{\text{in}})$ läßt sich durch Taylorreihenentwicklung näherungsweise mit

$$i_{\text{s}} \approx a_1 p_{\text{in}} + a_2 p_{\text{in}}^2 \quad (17)$$

darstellen. Die Klirrdämpfung der ersten Oberwelle $a_{\text{k}\,2}$ ergibt sich damit zu

$$a_{\text{k}\,2} \approx 20 \lg \left| \frac{2}{\hat{P}} \left(4 P_0 - \frac{n U_0}{R_{\text{e}} S M_0 (M_0 - 1)} \right) \right|. \quad (18)$$

Bild 4 zeigt die Klirrdämpfung $a_{\text{k}\,2}$ für Multiplikationsfaktoren $M_0 = 3$ bis 10, wobei als Parameter der dynamische Eingangswiderstand R_{e} des Empfängers gewählt wurde, da dieser sich in weiten Grenzen (Transimpedanzempfänger: R_{e} sehr klein, Hochimedanzempfänger: R_{e} sehr groß) ändern kann. Es ist deutlich erkennbar, daß R_{e} den wesentlichsten Einflußfaktor bildet. Zur Berechnung wurden für $\hat{P} = 8$ µW und für $P_0 = 16$ µW eingesetzt.

3. Zusammenfassung

Die in der Praxis auftretenden Frequenzmultiplexsignale sind in ihrer analytischen Beschreibung wesentlich komplizierter als das in diesem Beitrag angenommene Beispiel. Aufbauend auf den bisherigen Betrachtungen können hier lediglich einige Schätzungen gemacht werden:

- Bei Transimpedanzempfängern sind aufgrund des zwar frequenzabhängigen, aber im betrachteten Frequenzbereich noch sehr kleinen dynamischen Eingangswiderstands ($< 100\ \Omega$) entsprechend Bild 4 keine nichtlinearen Verzerrungen durch die dynamische Multiplikationsfaktorverschiebung zu erwarten.
- Bei Hochimpedanzempfängern ist bereits bei Eingangswiderständen $\gtrsim 20$ kΩ mit nichtlinearen Verzerrungen zu rechnen, sofern Multiplikationsfaktoren größer als 3 eingestellt werden. Nach Bild 4 kann die Klirrdämpfung $a_{\text{k}\,2}$ beispielsweise für $R_{\text{e}} = 100$ kΩ und $M = 3$ bis 10 mit $a_{\text{k}\,2} \approx 40$ dB geschätzt werden. Das bedeutet u.a., daß bei einem Frequenzmultiplexsignal ein im Frequenzbereich der ersten Oberwellen eines Übertragungskanals liegender wei-

Fortsetzung auf Seite 149

Das Rauschverhalten von Digital-Analog-Wandlern

W. Denda, KDT, Frankfurt (Oder)

Mitteilung aus der Akademie der Wissenschaften der DDR, Institut für Halbleiterphysik

In vielen Fällen wird mit Hilfe eines Digital-Analog-Wandlers (DAU) eine hochkonstante, programmierbare Bezugsspannung bereitgestellt. Dabei stört oft das der Ausgangsspannung überlagerte Rauschen, insbesondere das Funkelrauschen, das sich bei extrem niedrigen Frequenzen als zufällige Drift bemerkbar macht. Der Schaltungsentwickler muß dann Mittel und Wege suchen, dieses Rauschen zu minimieren, falls das wegen der existierenden Grenzen überhaupt möglich ist.

Es wird das Rauschen des DAU untersucht, insbesondere der Beitrag der einzelnen Quellen zum Gesamtrauschen. Testobjekte sind dabei die 12-bit-DAUs C 565 (HFO), MDAC 565 (Tesla) und AD 565 (Analog Devices). Die Aussagen können wegen der stets ähnlichen inneren Struktur der DAUs verallgemeinert werden, so daß sie auch für andere DAU-Typen gelten. Darüber hinaus ist ein Vergleich der 3 kompatiblen DAUs bezüglich ihrer Rauscheigenschaften möglich.

1. Rauschanalyse

Im Bild 1 ist das vereinfachte Rauschersatzschaltbild eines DAU dargestellt. Die wesentlichen Rauschquellen sind die Referenzquelle, der Pufferverstärker, die Stromquellen und der I-U-Wandler. Sie sind im Bild 1 durch entsprechende Spannungs- bzw. Stromquellen gekennzeichnet. Zur Ermittlung der Ausgangs-Rauschspannung U_0 ist es notwendig, den Einfluß der einzelnen Rauschquellen auf die Ausgangsspannung zu untersuchen.

Geht man davon aus, daß im 10-V-Bereich (unipolar) die Ausgangsspannung U_{OUT} etwa gleich der Referenzspannung U_{REF} ist, wenn alle Bits gesetzt sind, dann läßt sich der durch die Bits beeinflußbare Abbildungsfaktor k definieren zu

$$k = \frac{U_{OUT}}{U_{REF}}. \qquad (1)$$

Bezeichnet man mit $n = 0$ das LSB und mit $n = 11$ das MSB, dann gilt für ein einzelnes gesetztes Bit

$$k = 2^{(n-12)}. \qquad (2)$$

Es ist aus der Anordnung im Bild 1 unschwer zu erkennen, daß die Rauschspannung U_W des I-U-Wandlers direkt in die Ausgangs-Rauschspannung U_O eingeht, während bei den Rauschspannungen U_R (Referenzquelle) und U_P (Pufferverstärker) der Abbildungsfaktor k zu berücksichtigen ist. Bei dem Rauschstrom U_S/R_2 der Stromquellen muß man beachten, daß in erster Näherung das Quadrat des generierten Rauschstromes proportional dem fließenden Gleichstrom ist. Daraus folgt, daß sich der Rauschstrom U_S/R_2 mit der Wurzel aus k ändert.

Die einzelnen Rauschanteile sind unkorreliert, da sie in verschiedenen räumlichen Regionen entstehen. Damit ergibt sich das Quadrat der Ausgangs-Rauschspannung U_0 als Summe der Quadrate der Rauschanteile:

$$U_O^2 = (U_R^2 + U_P^2)\, k^2 + U_S^2\, k + U_W^2 \qquad (3)$$

(3) beschreibt das Rauschverhalten eines DAU vollständig und ist wegen ihrer Einfachheit für praktische Überlegungen gut anwendbar.

An dieser Stelle sind zwei Bemerkungen notwendig:

- Obwohl Bild 1 das „Innenleben" und die Peripherie eines C 565 darstellt, lassen sich die Betrachtungen dieses Abschnitts auch sinngemäß auf andere DAU-Typen übertragen. Für die grundsätzlichen Betrachtungen ist es unerheblich, wie viele Bits vorhanden sind bzw. ob die Referenzquelle und der I/U-Wandler integriert oder extern angeordnet sind. Auch läßt sich stets ein Abbildungsfaktor k nach (1) bestimmen, wobei natürlich (2) entsprechend aktualisiert werden muß.
- (3) gilt nicht nur für die breitbandig ermittelten Effektivwerte des Rauschens (gleiche Bandbreite ist Voraussetzung!), sondern auch für die spektrale Leistungsdichte W, da zwischen dem Quadrat des Effektivwertes und der spektralen Leistungsdichte ein linearer Zusammenhang besteht.

Um die Verhältnisse nach (3) zu illustrieren, ist im Bild 2 (Kurve a) die Ausgangs-Rauschspannung U_O eines C 565 in Abhängigkeit vom gesetzten Bit n dargestellt (wegen $n = \mathrm{ld}\,(k) + 12$ bedeutet $n = 12$,

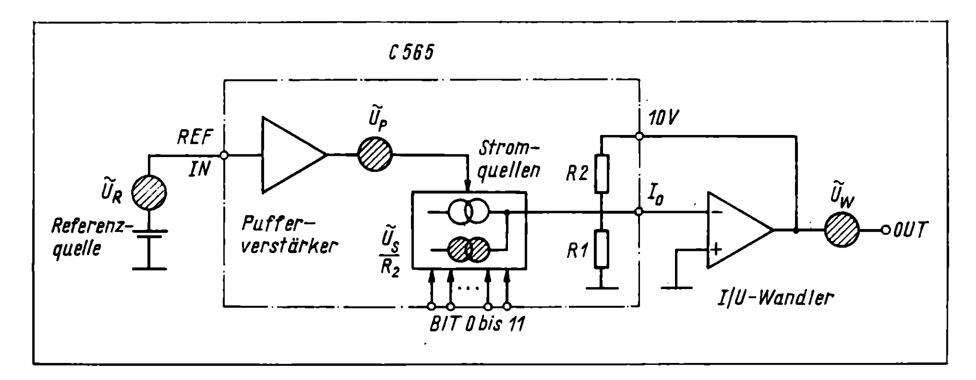


Bild 1. Vereinfachtes Rauschersatzschaltbild eines DAU
(Zwischen $R2$, $R1$, I_0 und den Stromquellen ist der Verbindungspunkt zu ergänzen!)

Fortsetzung von Seite 148

terer Übertragungskanal durch die in der APD entstehenden Oberwellen des 1. Kanals mit einer Störamplitude von 1% bez. der Nutzsignalamplitude im 1. Kanal gestört wird. Bei gleichen Bandbreiten und Nutzsignalamplituden beider Kanäle ergibt sich somit von vornherein ein Signal-Rausch-Abstand von nur 40 dB für den 2. Übertragungskanal. Abhilfe könnten z. B. schaltungstechnische Maßnahmen schaffen, die die Spannung über der APD konstanthalten.

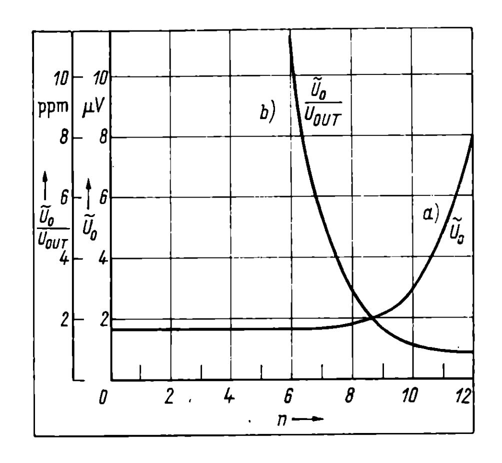
Denkbar wäre eine Stromteilerschaltung nach Bild 5, die mit dem unverstärkten, nichtentzerrten Eingangssignal des Hochimpedanzempfängers als Gegenkopplungssignal arbeitet.

NaA 47

Literatur

[1] *Smith, R. C.; Personick, S. D.:* Receiver Design for digital Fiber Optic Communication Systems. Part I & II. Bell Syst. Techn. Journal, 52 (1973) 6, S. 843–886
[2] *McIntyre, R. J.:* Multiplication Noise in Uniform Avalanche Diodes. IEEE Trans. Electron. Dev., ED 13 (1966) 1, S. 164–168
[3] *Smith, D. R.* u. a.: Receivers for Optical Fibre Communication Systems. Telecomm. Journal, 48 (1981) 11, S. 682
[4] *Grau, G.:* Optische Nachrichtentechnik. Eine Einführung. Berlin: Springer-Verlag 1981, S. 199–203

Dipl.-Ing. *Diethard Rau*, Deutsche Post, Rundfunktechnisches Zentralamt Agastr., Berlin, 1199



Tafel 1. Berechnete Werte für U_P und U_S

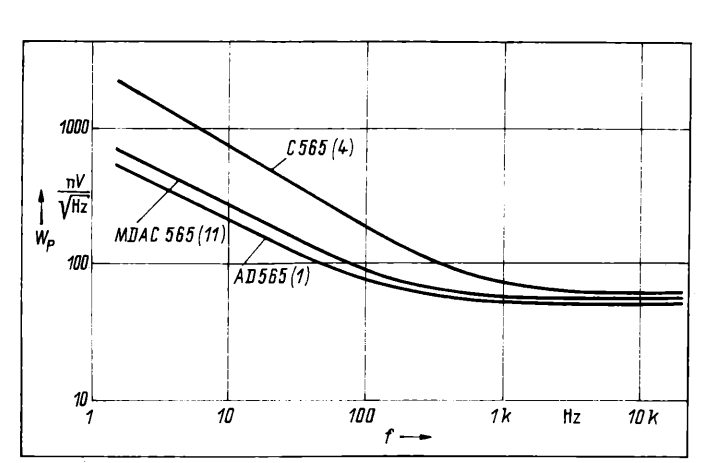
DAU	U_P μV	U_S μV
C 565 (4)	7,65	2,30
MDAC 565 (11)	6,56	1,96
AD 565 AJD (1)	5,65	1,91

Tafel 2. Gleichspannungen und Rauschspannungen verschiedener Referenzquellen

Quelle	U_{REF} V	U_R μV
C 565 (4)	10,0	18,5
MDAC 565 (11)	10,0	13,8
AD 565 AJD (1)	10,0	11,3
B 581 (20) (U = 15 V)	10,0	53,3
B 589 (6) (I = 0,5 mA)	1,25	17,3
2 × SZX 19/5,1 (I = 5 mA)	10,5	1,7

Bild 2. C 565 mit extrem rauscharmer Beschaltung: Ausgangs-Rauschspannung (a) und Verhältnis Rauschspannung/Gleichspannung in Abhängigkeit vom gesetzten Bit

Bild 3. Spektrale Rauschleistungsdichte des Pufferverstärkers ▼



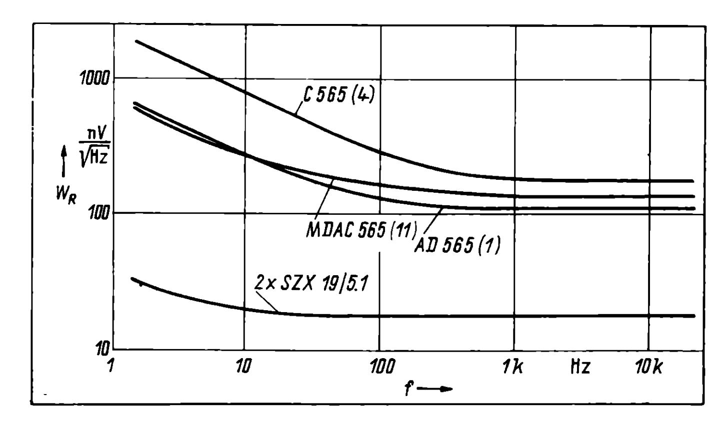


Bild 5. Spektrale Rauschleistungsdichte verschiedener Referenzquellen

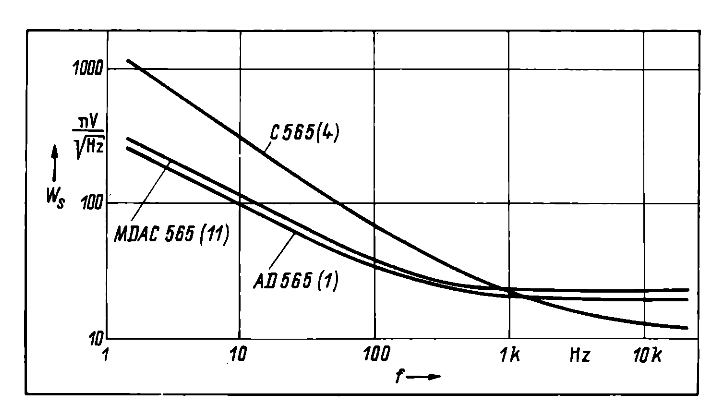


Bild 4. Spektrale Rauschleistungsdichte der Stromquellen

daß alle Bits gesetzt sind). Zur Ermittlung der Kurve wurden die äußeren Rauschquellen U_R und U_W extrem niedrig gewählt. Man erkennt deutlich, daß bei kleinen n-Werten der Kurvenverlauf durch U_W bestimmt wird. Bild 2b zeigt den Verlauf des Verhältnisses der Ausgangs-Rauschspannung U_O zur Ausgangs-Gleichspannung U_{OUT}. Die Kurve verdeutlicht, daß dieses Verhältnis nur für große n etwa konstant ist und zu kleinen n hin steil ansteigt (für $n = 0$ etwa 700 ppm).

2. Innere Rauschquellen

Die inneren Rauschquellen U_P und U_S sind durch den Anwender nicht beeinflußbar. Es ist aber günstig, die Größe der Rauschspannungen zu kennen, um die minimal mögliche Rauschspannung am Ausgang abschätzen zu können.
Die Bestimmung von U_P und U_S erfolgt auf der Grundlage von (3). Wenn man die Größe von U_W nicht direkt ermittelt hat (siehe Abschn. 4), nutzt man die Tatsache aus, daß für kleine n mit hinreichender Genauigkeit $U_O = U_W$ ist (siehe Bild 2). Die Rauschspannung U_R der Referenzquelle kann durch direkte Messung bestimmt werden. Damit bleiben in (3) zwei unbekannte Größen (U_P und U_S) übrig. Um diese zu berechnen, sind zwei Messungen der Ausgangs-Rauschspannung U_O bei unterschiedlichem n erforderlich. Aus dem Bild 2 erkennt man, daß zweckmäßigerweise dazu das MSB ($n = 11$) und ein etwas kleineres Bit ($n = 9$) benutzt werden sollte.
Unter Verwendung von (2) und (3) erhält man dann

$$U_P^2 = \frac{U_{O\,11}^2 \cdot k_9 - U_{O\,9}^2 \cdot k_{11} - U_R^2\,(k_{11}^2 \cdot k_9 - k_{11} \cdot k_9^2) + U_W^2\,(k_{11} - k_9)}{k_{11}^2 \cdot k_9 - k_{11} \cdot k_9^2} \quad (4)$$

und

$$U_S^2 = \frac{U_{O\,11}^2 - (U_R^2 + U_P^2) \cdot k_{11}^2 - U_W^2}{k_{11}}. \quad (5)$$

Die Tafel 1 gibt Auskunft über die ermittelten Effektivwerte (Bandbreite: 10 Hz bis 10 kHz) der verschiedenen DAUs.
Aus gemessenen Rauschspektren der Referenzquelle, des I-U-Wandlers und des Ausgangs ($n = 11$ und $n = 9$) wurden die Rauschspektren des Pufferverstärkers und der Stromquellen punktweise unter Benutzung von (4) und (5) berechnet. Sie sind in den Bildern 3 und 4 dargestellt.

3. Referenzquelle

Unter dem Aspekt, daß der DAU am Ausgang eine möglichst geringe Rauschspannung aufweisen soll, ist es nicht sinnvoll, die interne Referenzquelle des DAU zu benutzen. Als alternative Quellen bieten sich die Referenzquellen B 581 und B 589 sowie ausgesuchte Z-Dioden an. In Tafel 2 sind die stabilisierten Gleichspannungen und die überlagerten Rauschspannungen (Bandbreite: 10 Hz bis 10 kHz) einiger Referenzquellen aufgeführt.
Die Referenzquelle B 581 scheidet wegen ihrer sehr großen Rauschspannung aus. Bei der Referenzquelle B 589 muß man beachten, daß die Gleichspannung um den Faktor 8 verstärkt werden muß, um die geforderte Spannung von 10,0 V zu erreichen. Dadurch erhöht sich die Rauschspannung auf 138,4 μV. Besser wäre es, 8 Stück B 589 in Reihe zu schalten, da sich in diesem Falle wegen der Addition der Rauschleistungen nur eine resultierende Rauschspannung von 48,9 μV ergeben würde. Aber auch diese ist, gemessen an der Rauschspannung der DAUs, noch viel zu groß.
Extrem niedrige Rauschspannungen zeigen die Z-Dioden SZX 19/5,1 und SZX 21/5,1. Von 30 Dioden aus unterschiedlichen Produktionszeiträumen waren 24 Dioden brauchbar ($U_R < 1,3$ μV). Bei

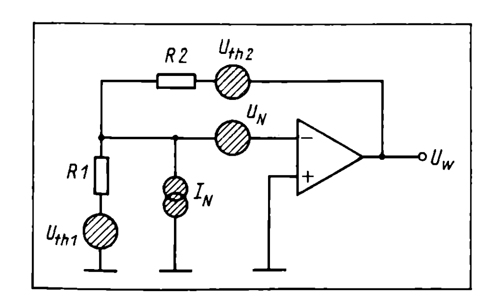


Bild 6. Rauschquellen des I/U-Wandlers

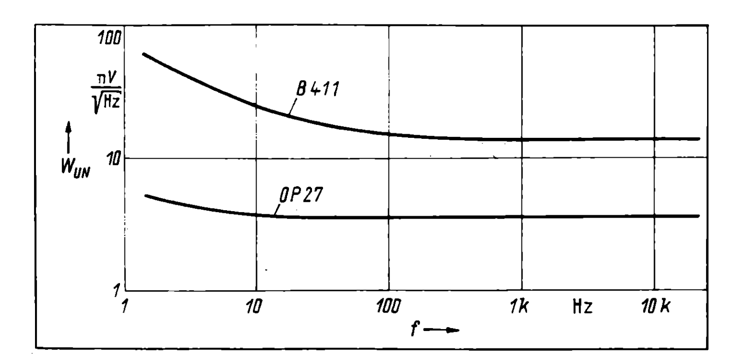


Bild 7. Spektrale Leistungsdichte der eingangsbezogenen Rauschspannung U_N

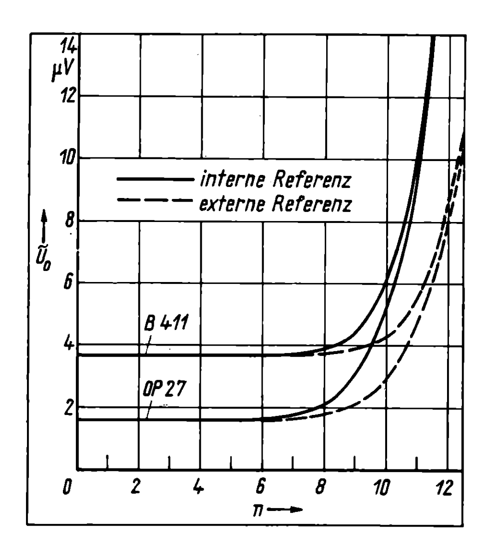


Bild 8. C 565: Ausgangs-Rauschspannung in Abhängigkeit vom gesetzten Bit

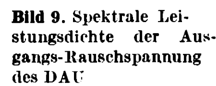
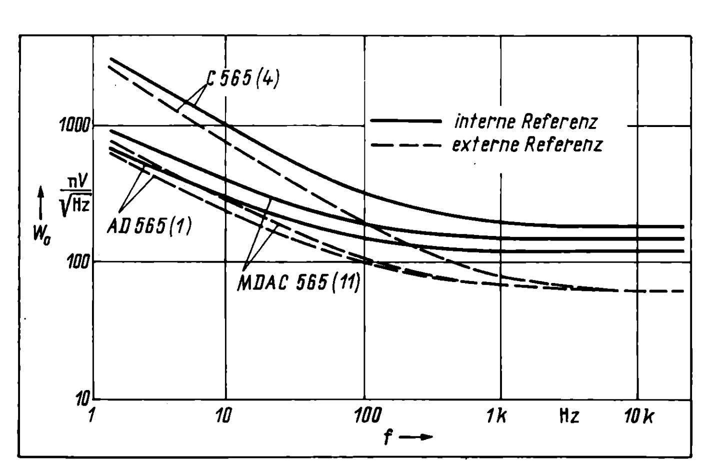


Bild 9. Spektrale Leistungsdichte der Ausgangs-Rauschspannung des DAU



den restlichen Dioden wurde starkes Funkel- bzw. Stoßrauschen festgestellt. Wegen dieser niedrigen Rauschspannung eignet sich eine Reihenschaltung von 2 Z-Dioden ausgesprochen gut als Referenzquelle. Die etwas zu hohe Referenzspannung kann durch einen zusätzlichen Widerstand vor dem Referenzeingang des DAU (etwa 1 kΩ) ausgeglichen werden. Nachteilig wirkt sich der relativ hohe Temperaturkoeffizient der Z-Dioden aus. Abhilfe kann hier ein elektrisch beheizter Mini-Thermostat schaffen, der direkt auf der Leiterplatte angeordnet ist.

Im Bild 5 sind die spektralen Leistungsdichten der untersuchten DAUs (typische Vertreter) sowie einer Reihenschaltung von 2 Z-Dioden SZX 19/5,1 dargestellt. Neben einem niedrigeren Absolutpegel des Rauschens zeigen die Z-Dioden im Gegensatz zu den DAUs auch einen wesentlich geringeren Funkelrauschanteil.

4. I-U-Wandler

Der I-U-Wandler legt den Rauschpegel bei den niederwertigen Bits fest. Die Betriebsart des OPV (Operationsverstärker) ist aus dem Bild 1 zu erkennen. Im Bild 6 sind für diese Betriebsart die einzelnen Rauschquellen extrahiert dargestellt, d.h., die Widerstände R_1 und R_2 sowie der OPV sind rauschfrei. Danach ergibt sich die Rauschspannung U_W am Ausgang des OPV (vgl. Bild 1) zu

$$U_W{}^2 = U_{th\,2}{}^2 + U_{th\,1}{}^2 \left(\frac{R_2}{R_1}\right)^2 + U_N{}^2 \left(\frac{R_1 + R_2}{R_1}\right)^2 + I_n{}^2 R_2{}^2 \qquad (6)$$

mit

$$U_{th}{}^2 = 4kTBR. \qquad (7)$$

Die durch den DAU vorgegebenen Widerstandswerte liegen für den C 565 und die dazu kompatiblen Typen bei $R_1 = 4{,}0\,\text{k}\Omega$ und $R_2 = 5{,}0\,\text{k}\Omega$. Damit sind durch den Anwender nur noch die eingangsbezogenen Rauschgrößen U_N und I_N des OPV durch die Auswahl eines geeigneten OPV zu beeinflussen. Bei der Auswahl des OPV ist natürlich zu beachten, daß der Eingangsstrom des OPV möglichst weit unterhalb eines LSB (etwa 500 nA) liegt.

Als Testobjekte wurden ein B 411 (HFO) mit $U_N = 1400$ nV und $I_N = 5{,}8$ pA sowie ein OP 27 (PMI) mit $U_N = 360$ nV und $I_N = 44$ pA (Bandbreite: 10 Hz bis 10 kHz) benutzt. Im Bild 7 sind für diese beiden OPV die spektralen Leistungsdichten W_{UN} der eingangsbezogenen Rauschspannung U_N dargestellt.

5. Ausgangs-Rauschspannung

In den vorangegangenen Abschnitten wurde aufgezeigt, wie sich die Ausgangs-Rauschspannung zusammensetzt und welche Quellen man beeinflussen kann. Zur Illustration des dargestellten Sachverhalts dienen die Bilder 8 und 9.

Im Bild 8 ist für den C 565 die Ausgangs-Rauschspannung in Abhängigkeit vom gesetzten Bit aufgetragen (Bandbreite: 10 Hz bis 10 kHz). Als Referenzquelle wurden die interne Quelle und die Reihenschaltung von 2 Z-Dioden SZX 19/5,1 (extern) benutzt. Als I-U-Wandler kamen die OPV-Typen B 411 und OP 27 zum Einsatz. Man erkennt deutlich, daß entsprechend (3) die Referenzquelle die oberen Bits und der OPV die unteren Bits beeinflußt.

Im Bild 9 ist die spektrale Leitungsdichte der Ausgangs-Rauschspannung der untersuchten DAUs aufgetragen (alle Bits gesetzt, $n = 12$). Als Parameter wurde wie im Bild 8 die Referenzquelle benutzt. Der Einfluß des als I-U-Wandler benutzten OPV ist in dieser Darstellung ($n = 12$!) nicht zu erkennen. Die dargestellten Kurvenverläufe zeigen sehr deutlich, daß bei niedrigen Frequenzen die inneren Rauschquellen U_P und U_S (große Funkelrauschanteile) dominieren, während sich bei hohen Frequenzen der Einfluß der Referenzquelle stärker durchsetzt.

6. Zusammenfassung

Die innere Struktur eines DAU wurde analysiert und daraus die wesentlichen Rauschquellen ermittelt. Der Beitrag der einzelnen Rauschquellen zur Ausgangs-Rauschspannung wurde definiert. Die gewonnenen Ergebnisse gelten gleichermaßen für die integrale und die spektrale Darstellung des Rauschens. Durch die Wahl der äußeren (nicht integrierten) Baugruppen des DAU kann der Anwender die Größe der Ausgangs-Rauschspannung wesentlich beeinflussen. Die dargestellten Meßergebnisse ermöglichen einen Vergleich der DAUs verschiedener Hersteller.

NaA 26

Dr.-Ing. *Wolfgang Denda*, Akademie der Wissenschaften der DDR, Institut für Halbleiterphysik, Walter-Korsing-Str. 2, Frankfurt (Oder), 1200

Stochastische Modellierung von fehlertoleranten Systemen

I. Jozwiak, Wroclaw; U. Kosmol, Dresden

Mitteilung aus dem Polytechnikum Wroclaw und dem VEB Robotron-Elektronik Dresden

Ein Prinzip zur Modellierung von fehlertoleranten Systemen mit gesteuerten stochastischen Prozessen wird erläutert. Es werden Real-Time-Multicomputer-Systeme betrachtet. Grundlage ist die Dekomposition in Teilprozesse der Zuverlässigkeitsänderungen, Systemsteuerung und Task-Realisierung. Die Analyse läßt Aussagen zu funktionellen Kenngrößen und zur optimalen Steuerung zu.

1. Allgemeine Grundlagen und Zielstellung

Im wesentlichen werden 3 grundlegende Verfahren bei der Modellierung fehlertoleranter Computersysteme angewandt:

- experimentelle Methode

  Untersuchung zur Realisierbarkeit neuartiger Architekturen mit der entsprechenden Software, hoher Kostenaufwand und geringe Vergleichsmöglichkeit mit anderen Systemen, bedingt durch die Applikationsbezogenheit des Systems
- Simulationsmethode

  genaues Systemverhalten bei hohem Programmieraufwand analysierbar; Änderungen im System erfordern oft wesentliche Änderungen im Simulationsprogrammpaket
- analytische Methode

  hoher Abstraktionsgrad; Diskrepanz zwischen Modell und Wirklichkeit aufgrund mathematischer Beschränkungen; Schwierigkeiten bei der Erfassung der Eingangsparameter und Implementierung von Softwareerscheinungen.

Selten erfaßt eine Methode die Gesamterscheinungen des modellierten Systems. Das vorzustellende Prinzip ist den analytischen Methoden zuzuordnen, wobei die Diskrepanzen zwischen Wirklichkeit und Modell gering gehalten werden.
Das Modell sollte folgenden Forderungen genügen:

- analytisches Modell für eine breite Klasse von Systemen mit wählbarem Abstraktionsniveau
- dem Projektanten zugängliche Eingangsparameter
- Möglichkeit einer automatisierten Herstellung zur Implementation in CAD-Systeme
- Ableitung von funktionsbezogenen Bewertungsgrößen.

Fehlertoleranz ist die Fähigkeit eines Systems, seine spezifizierten Funktionen bei Vorhandensein von bestimmten Fehlern weiterhin auszuführen, ohne es unbedingt dem Nutzer direkt zu signalisieren. Die notwendige Voraussetzung für diese Fehlertoleranz ist die gezielte Anwendung von Redundanz in Hardware und Software zur Fehlermaskierung in Verbindung mit entsprechenden Steueralgorithmen.
Daraus folgt, daß für die Fehlertoleranz stets das System in seiner Gesamtheit zu betrachten ist. Reine Redundanzuntersuchungen sind unzureichend. Günstige Voraussetzungen sind implizit in Multicomputer-Systemen mit modularem Aufbau und Multi-Task-Verarbeitung anzutreffen. Generell kann man diese in zwei Klassen unterteilen:

General Purpose System	Real Time System
funktionelles Bewertungskriterium:	
— Durchsatz	— Reaktionszeit
Programmstruktur:	
— veränderlich über die Betriebszeit	— relativ starr während der Betriebszeit
— Nutzer-Task-Reihenfolge weniger bedeutend	— Applikations-Task-Reihenfolge wesentlich
Steuerung:	
— a priori nicht festlegbar	— a priori festlegbar (ein Anwendungsfall)
Fehler:	
— selten drastische Folgen, Behebung in Wartungspausen	— können Gefahrensituationen bewirken, schnelle Reaktion erforderlich.

Die nachfolgenden Betrachtungen beziehen sich auf Real-Time-Multicomputer-Systeme mit a priori festgelegter Reihenfolge der Taskabarbeitung.
Es wird angenommen, daß zur Erzeugung einer Reaktion eine gewisse Taskserie abgearbeitet werden muß. Bei der Beurteilung der Fehlertoleranz ist die Fehlerbestimmung wichtig. Im System ist ein Fehler präsent, sofern definierte Abarbeitungszeiten einer Taskserie (Deadlines) überschritten werden. Ein Hardwarefehler ist nicht unbedingt gleichbedeutend mit einem Systemfehler.

2. Modell

Ein Real-Time-Multicomputer-System mit Multi-Task-Verarbeitung ist gekennzeichnet durch:

a) die Hardware-Struktur in Form einer endlichen Menge C von Ressourcen, die untereinander nach einer Relation N verbunden sind und einem Fehler- und Erneuerungsprozeß unterliegen. Dieser Prozeß wird mit dem Zufallsvektor R_t als Kennzeichnung für die Hardwarezuverlässigkeit belegt,

Beispiel: $R_{t1} = (1, 1, 1)$ $\quad$ $R_{t2} = (1, 1, 0)$

$\boxed{C_1} \overset{N_{12}}{—} \boxed{C_2}$ $\quad$ $\boxed{C_1}$ $\boxed{C_2}$ N_{12} — defekt

b) die Applikationsstruktur in Form einer endlichen Menge Z, von den auf den Ressourcen zu realisierten Task, die entsprechend dem Datenfluß und dem Startzeitpunkt geordnet sind,

Beispiel: $\boxed{Z_1} \rightarrow \boxed{Z_4}$

$\boxed{Z_2} \nearrow$

c) die Tasksteuerungsstruktur in Form eines Task-Queueing-Allocation-Algorithmus QA, der die Applikationsstruktur um entsprechende Systemtasks (Kommunikation zwischen Computern, Rekonfiguration) ergänzt und zu jedem Zeitpunkt für jede Hardware-Struktur R_t die nächsten Task einsteuert, dargestellt als aktuelle Realisierung QA_{Rt},

Beispiel: QA_{Rt1} $\quad$ QA_{Rt2}

$Z_{11} \rightarrow Z_3 \rightarrow Z_{41}$ $\quad$ $Z_{11} \rightarrow Z_{21} \rightarrow Z_{41}$

$\nearrow$ $\vdots$

Z_{22} $\quad$ Comm. $Z_2 \rightarrow Z_1$ $\quad$ $\widehat{Z_{22}}$.

Die Aufgabe des Systems besteht in der zeitgemäßen Realisierung einer maximalen Anzahl von Applikationstasks. Dies bedingt die Einführung eines Vektors E_t, der zum Zeitpunkt t realisierten Tasks, mit dem zeitliche Kenngrößen abgeleitet werden können. Im System sind die Vektoren R_t und E_t jederzeit beobachtbar. In Abhängigkeit von deren aktuellem Wert leitet das System eine der folgenden Entscheidungen ab:

1. Ausführung der nächsten Task infolge Änderung von E_t ohne Änderung von R_t.
2. Änderung von QA_{Rt} infolge Änderung von R_t (bei Mehrfachredundanz wird eine Änderung erst eingeleitet, sobald die Redundanz wegfällt — erste Etappe einer Fehlertolerierung).
3. Systemstop infolge eines System-Breakdown.

3. Systemverhalten als stochastischer Prozeß

Der Zustand des Systems ist jederzeit eindeutig bestimmbar als

$S = (R_t, E_t, QA_t)$

$QA_t = QA_{Rt}$.

Die Zustandsübergänge erfolgen durch

1. Defekt einer Ressource
2. Reparatur einer Ressource
3. Abschluß einer Task.

In einem fehlerfreien System ist $R_t = \text{const}$ und

$f = QA\ (R_t) = QA$ sowie

$E_{t+1} = QA\ (E_t)$

in einem fehlerfreien System ist $R_t \neq \text{const}$ und

$f = QA\ (R_t)$ sowie

$E_{t+1} = (QA\ (R_t))\ (E_t)$.

Die Änderungen von R_t erfolgen stochastisch, somit werden auch $QA\ (R_t)$ und demzufolge E_t stochastisch geändert. Die Analyse der Abhängigkeiten läßt eine neuartige Klassifikation zu.

■ Definition:

Es sei (D_t), teT ein stochastischer Prozeß, dessen Realisierungen zu teT Elemente einer nichtleeren Entscheidungsmenge D sind. Dann ist der Prozeß $(X_t\ {}^{D}t)$, teT, DeT ein gesteuerter stochastischer Prozeß, dessen Steuerung durch D_t, teT bestimmt ist [1].

Es sei $D_t = R_t$, $X_t = QA_t \rightarrow (QA_t\ {}^{R}t)$,

dann ist $(QA_t\ {}^{R}t)$ ein gesteuerter stochastischer Prozeß, der besagt, daß der aktuelle Steuerungsalgorithmus von dem durch die Zuverlässigkeitsänderungen beschriebenen stochastischen Prozeß festgelegt wird.

Es sei $D_t := QA_t$, $X_t = E_t \rightarrow (E_t\ {}^{QA}t)$.

Der Taskrealisierungsprozeß wiederum wird von der aktuellen Realisierung QA bestimmt.
Das Gesamtverhalten wird endgültig als

$$\left({}_{E_t} \quad QA_t \quad {}^{R_t} \right)$$

formuliert und stellt einen zweistufig gesteuerten stochastischen Prozeß dar.
Eine derartige Formulierung erlaubt genau festzustellen, welche Aufgabe, wann und wo im System realisiert wird. Dies ist Voraussetzung für die Ermittlung der Reaktionszeiten, die an die Realisierung einer Taskserie gebunden sind. Bei Einführung einer Kostenvariablen [2] läßt sich noch die Wichtigkeit einer Task feststellen und ein fail-soft Verhalten nachbilden. Mit dieser Herangehensweise ist eine detailliertere Untersuchung als in entsprechenden Bedienmodellen möglich.
Zusätzlich werden von der Kostenvariablen Aussagen zur Optimalität der Steuerung, d.h. des QA-Algorithmus, möglich. Die Verknüpfung der Wichtigkeit von Task mit den in der gegebenen Konfiguration erzielten Reaktionszeiten läßt Rückschlüsse auf eine zweite Art der Fehlertoleranz zu, indem ermittelt wird, bei welcher Fehlermenge und welchem QA-Algorithmus die Task-Deadlines überschritten werden.

4. Probleme bei der Bestimmung des Entscheidungsprozesses

Ein wesentliches Problem bei der Modellierung ist die stochastische Nachbildung des realisierten Entscheidungsprozesses $(QA\ {}^{R}t)$, teT. Die Eingangswerte sind durch die Wichtigkeit der Task, die in der aktuellen Hardware-Struktur (definiert durch R_t) zu realisieren sind und die Hardware-Struktur selbst bestimmt. Das Ergebnis stellt einen Algorithmus $QA\ (R_t)$ dar, dessen einzelne Realisierung QA_t die Abarbeitsfolge der Tasks auf den Ressourcen bilden.
Im Real-Time-System erfolgt die Änderung $QA\ (R_t)$ zu $QA\ (R_{t+1})$ stochastisch, während der Übergang von QA_t zu QA_{t+1} deterministisch stattfindet. Diese Erscheinung konnte noch nicht als einheitlicher stochastischer Prozeß gefaßt werden.

Zusätzliche Schwierigkeiten treten auf bei der

- Berücksichtigung asynchroner Ereignisse (event-driven task)
- realitätsbezogener Nachbildung von Recovery-Prinzipien.

5. Schlußfolgerungen

Das skizzierte Prinzip zeigt, daß die Steuerungsproblematik bei gelösten Redundanzfragen für das Erreichen von Fehlertoleranz von immenser Bedeutung ist. In der gewählten Betrachtungsweise fand eine Transformation der Fehlertoleranz von der Hardware ins Gebiet der stochastischen Steuerungstheorie statt. Redundanz in Hardware und Software definiert die Menge der tolerierten Fehler. Erst bei Einbindung in die Steuerung wird eine Aussage über funktionsorientierte Fehlertoleranz möglich.

Die vorgestellte Methode ist gekennzeichnet durch

- beliebig wählbares Modellierungsniveau
- relative Unabhängigkeit von der Netzwerkstopologie
- dem Projektanten zugängliche Eingangsparameter (Fehlerraten, Reparaturraten, Taskverarbeitungszeiten, Abarbeitungsalgorithmen)
- automatisierte Modellerstellung
- implizite Konstruktionsmethodologie (Hardware — Taskstruktur — Steuerung).

Sie sind anwendbar für

- dynamische Analyse noch nicht existenter Systeme
- Optimierung der Steuerungsalgorithmen
- Schwachstellenanalyse
- Angabe quantitativer und qualitativer Fehlertoleranzkenngrößen zum Systemvergleich.

NaA 27

Literatur

[1] *Müller, P. H.* (ed.): Lexikon der Stochastik. Berlin: Akademie-Verlag 1983
[2] *Kosmol, U.:* Stochastik Modelling of Fault Tolerant Computing Systems Considering Task Queueing. Proc. Micro-Computer '86, Wydawnictwo PWr, Wrocław 1986, s. 139–144, Scientific Papers of the Institute of Engineering Cybernetics of the Technical University of Wrocław, No. 73

Dr.-Ing. *Ireneusz Jozwiak*, Technische Universität Wrocław, Institut für Kybernetik, ul. Janiszewskiego 11–17, Wrocław, PL–50–372

Dipl.-Ing. *Uwe Kosmol*, VEB ROBOTRON Elektronik Dresden, Abt. E 92, Leningrader Str. 15, Dresden, 8010

Kurzinformationen

• Die japanische Fernmeldegesellschaft NTT hat im April ihren **ISDN-Dienst** (mit der Bezeichnung INS NET 64) zunächst in Tokio, Osaka und Nagoya in Betrieb genommen. Die Netze werden in ihrer gegenwärtigen Form mit 64 kbit/s betrieben. Bis Ende nächsten Jahres soll ISDN „in allen größeren Städten des Landes" verfügbar sein. Daneben beginnt NTT im nächsten Jahr unabhängig von den bisherigen Ausbauplänen mit der Einrichtung von „INS NET 1500" einem Datennetz, das mit 150 kBit/Schaltgeschwindigkeit arbeiten kann. Diese Hochleistungssysteme dürften zunächst in Zentren mit hohem Datenverkehrsaufkommen versuchsweise in Betrieb genommen werden. Die japanischen ISDN-Netze sollen einige kundenfreundliche Spezialangebote erhalten. Dazu gehört u.a. die Identifikation des rufenden Teilnehmers. Der Empfänger erhält auf einem Monitor die Fernsprechnummer des Anrufers angezeigt. Ferner werden die entstehenden Gebühren dem rufenden ISDN-Teilnehmer unmittelbar nach Auftrennung einer Verbindung auf einem Monitorschirm dargestellt.

• In den USA, in Europa und in Japan verstärken sich gegenwärtig die Forschungs- und **Entwicklungs-Aktivitäten,** die **zu neuen Halbleiter-Elementen** führen sollen, bei denen der sogenannte Quanten-Effekt ausgenutzt werden soll. Die Nutzung dieses Effekts hängt von einem neuen Verständnis der Energieband-Lücken, der Valenz-Bänder und der bereits bekannten Quanten-„Schächte" (wells) ab. Bei einem geeigneten Aufbau gelangt man zu Kanälen im Halbleiter-Material und zu Strukturen, die unter Bezug auf die Elektronen-Bewegungen als „eindimensional" bezeichnet werden. Bekannt sind Vorarbeiten für Quanten-Effekt-Elemente u.a. bei Firmen und in Instituten der USA, Japans, der Niederlande und in der Bundesrepublik. Auslöser für das wachsende Interesse ist die Tatsache, daß man mit der heutigen Silizium-Halbleitertechnik nach allgemeiner Ansicht die Geometrie (Linien-Auflösung) von 0,25 µm auch langfristig nicht unterschreiten kann. Mit Quanteneffekt-Strukturen sollen nach heutiger Erwartung amerikanischer Forschungsstellen jedoch Geometrien von 0,01 µm möglich sein.

Fachtagungen

Mikroelektronik '88

Die 6. Fachtagung „Applikation Mikroelektronik" des Fachverbands Elektrotechnik der KDT war eine gemeinsame Veranstaltung der Ingenieurorganisationen sozialistischer Länder, die vom Fachverband Elektrotechnik unter dem Namen „Mikroelektronik '88" organisiert worden war. Sie fand vom 7. bis 9. Dezember 1988 in Dresden-Niedersedlitz statt. Mit der Tagung wurde auch Bilanz zur Realisierung des KDT-Initiativprogramms vom Mai 1986: „Entwicklung der Elektrotechnik/Elektronik unter den Bedingungen der intensiven Entwicklung und Applikation der Mikroelektronik" gezogen.

An der Schwelle der 90er Jahre ist die Mikroelektronik durch den Übergang zur höchstintegrierten Technik mit Abmessungen im Submikrometerbereich gekennzeichnet. Dadurch stellt die weitere Umsetzung der Beschlüsse des XI. Parteitages der SED neue höhere Anforderungen an Wissenschaftler, Ingenieure und Techniker in den Betrieben und wissenschaftlichen Einrichtungen. Den hohen Anforderungen der VLSI/ULSI-Technik — so stellte Prof. Dr. *Jähn*, Vorsitzender des Fachverbands Elektrotechnik der KDT, zur Eröffnung der Tagung fest — müssen sich die Ingenieurorganisationen in den sozialistischen Ländern stellen. Das erfordert vor allem auch eine Intensivierung der geistig-schöpferischen Arbeit. Mit dem praxisorientierten Informations- und Erfahrungsaustausch zu den bisher erreichten Ergebnissen und zu weiteren, geplanten Vorhaben zur Mikroelektronik sollte die Tagung einen maßgeblichen Beitrag leisten und die Teilnehmer der Tagung befähigen, daraus entscheidende Aktivitäten am eigenen Arbeitsplatz zur Erhöhung der Leistungskraft der Volkswirtschaft abzuleiten.

Diesem Anliegen trug das Tagungsprogramm mit über 100 Vorträgen von Referenten aus der DDR, UdSSR, VRB, VRP, ČSSR und UVR Rechnung, die in Plenarvorträgen und 7 Arbeitssektionen gehalten wurden. Etwa die Hälfte der Vorträge stellte Ergebnisse vor, die durch die Tätigkeit der Ingenieurorganisation in den befreundeten Ländern erzielt worden sind.

Hervorzuheben waren auch die Beratungs- und Informationsmöglichkeiten, auf der Fachtagung über

- das aktuelle Sortiment moderner elektronischer Bauelemente aus der DDR
- die neuesten Hardware-Lösungen bei Kleinrechnern (die Leistungsfähigkeit der Gerätesysteme konnte während der Tagung selbst getestet werden)
- die im Rahmen einer Software-Börse gespeicherten über 1500 Software-Lösungen
- die mehr als 20 nachnutzbaren mikroelektronischen Lösungen, die im Rahmen einer Poster-Diskussion vorgestellt wurden
- den aktuellen Stand der verfügbaren Informationsquellen aus dem staatlichen Informationsdienst Mikroelektronik der DDR.

Hervorzuheben war auch die hervorragend wirksame Simultan-Übersetzung (russisch, englisch, deutsch) bei dem hohen Anteil ausländischer Referenten in 2 der 3 jeweils parallel tätigen Arbeitssektionen.

Mit dem Hauptvortrag der Tagung „Von der Mikroelektronik zur Nanoelektronik" stellte Prof. Dr. *Köhler* (THI) die Entwicklungstendenzen und Anforderungen heraus, die mit dem Übergang zur Submikrometertechnik zu bewältigen sind. Strukturabmessungen kleiner 1 µm bis hinunter zu 0,5 µm stellen in den 90er Jahren hohe Anforderungen an Entwickler, Technologen und vor allem an die Hersteller der Spezialausrüstungen, um z.B. die Reinraumanforderungen der dann notwendigen Klasse 1 (d.h., in einem Volumen von 28 Litern darf höchstens noch 1 Staubpartikel von $\leqq$ 0,5 µm vorhanden sein) zu erfüllen und hinsichtlich des Übergangs der Strukturerzeugung von lichtoptischen Verfahren zu Strahltechniken. Nur so kann man Speicherschaltkreise mit 16, 64 bzw. sogar 256 Mbit Kapazität bei annehmbaren Ausbeuten herstellen.

Die perspektivischen Entwicklungsrichtungen und mögliche Leistungsfähigkeit der Mikroelektronik wurden von Prof. Dr. *Enderlein* und Prof. Dr. *Nowick* in den Plenarvorträgen „Halbleitermikrostrukturen: Supergitter-Multi-Quantum-Wells" und „Amorphe Halbleiter-Supergitter und -Vielfachschichten" dargestellt. Beide Vorträge zeigten, daß mit der Annäherung der mikroelektronischen Parameter an physikalische Grenzen und Naturkonstanten nicht nur der wissenschaftliche Aufwand wächst, sondern auch neue Anforderungen an die ingenieurtechnische Umsetzung der Forschungsergebnisse in die „elektronische Praxis" sichtbar werden.

Zwei weitere Plenarvorträge stellten die engen Beziehungen der Mikroelektronik zur Hauptanwendungsrichtung Computertechnik dar. Prof. Dr. *Cimander* und Prof. Dr. *Bernstein* informierten in ihren Vorträgen „Hochtechnologie und Computertechnik" und „Entwicklungstendenzen von Multi-Prozessor-Systemen" über die Leistungsfähigkeit mikroelektronischer Systeme bei der Informationsverarbeitung, zeigten aber auch die ökonomischen Erfordernisse auf, so wird z.B. die Entwicklung des 16-Mbit-DRAM etwa 3 Mrd. DM kosten.

Der Plenarvortrag von Dr. *Prischmann* „Stand beim unipolaren Schaltkreissortiment des VEB Kombinat Mikroelektronik" schloß den weit gespannten Bogen der Einführungsvorträge. Er informierte über das gegenwärtig verfügbare Bauelemente-Sortiment und die nächsten Aufgaben des Kombinats Mikroelektronik. Aus dem Vortrag ließ sich ableiten, daß die DDR seit der 5. Fachtagung „Applikation Mikroelektronik" 1986 einen beachtlichen Schritt nach vorn gemacht hat. Mit den 16-bit-Mikroprozessoren, den 64-kbit- und 256-kbit-Speichern, die in absehbarer Zeit durch 32-bit-Mikroprozessoren sowie 1-Mbit- und 4-Mbit-Speicher erweitert werden, hat die DDR ihren Platz unter den wenigen hochentwickelten Industrieländern gefestigt, die diese Technik beherrschen.

Komplex elektronische Bauelemente

In den 4 Arbeitssektionen standen

- Ergebnisse und Probleme des Entwurfs hochintegrierter Schaltkreise
- Informationen zu neu entwickelten Bauelementelinien
- Prüfung und Zuverlässigkeit elektronischer Bauelemente

im Mittelpunkt.

In den Vorträgen wurde u.a. informiert über:

- den erreichten Stand bei Mikroprozessoren und ihre Einsatzgebiete — speziell in der Computertechnik (DDR, VRB)
- das Sortiment von Wandler-Schaltkreisen und ihre Einsatzmöglichkeiten (DDR)
- die Entwicklung und Anwendung von SMD-Bauelementen
- die Verwendung von Chip-Carriers für die SMT (ČSSR)
- die in der DDR erkennbaren Sortimentsschwerpunkte für die Entwicklung und den Einsatz von anwendungsspezifischen Schaltkreisen (ASICs); unverkennbar ist dabei, daß auch die Kundenwunschschaltkreise des ISA-Systems vom Halbleiterwerk Frankfurt (Oder) für spezifische Anwendungsfälle ökonomisch bedeutungsvoll sind
- ein interessantes und wichtiges Anwendungsfeld der CMOS-Technik — die verfügbaren Codec-Schaltkreise und die Leistungsfähigkeit der CMOS-Decoder (VRB)
- das optoelektronische Bauelementesortiment der DDR und ihre Einsatzbreite (Si-Fotodioden, LED und LEA, Einsatz von CCD-Zeilen- und Matrixsensoren).

Komplex Layoutentwurf von VLSI-Bauelementen

Hier standen wissenschaftliche Arbeitsergebnisse zur Diskussion:

- Probleme und Lösungsmöglichkeiten zur „relativen und absoluten Plazierung"
- Verdrahtungsflächen-Strukturierung
- Programmsysteme zur Spannungs- und Masseverdrahtung
- Gesichtspunkte der Kanalverdrahtung.

Damit wurde ein geschlossener Überblick über die erforderlichen Arbeiten zum automatischen Layoutentwurf von VLSI-Bauelementen gegeben. Auch die Beiträge ausländischer Fachkollegen

- zum thermischen Verhalten von Halbleiterbauelementen (VRP/UVR)
- zum praktikablen Modell „STOLEM — ein Gate-Array-System" (VRP)

machten mit neuen wissenschaftlichen Untersuchungsergebnissen bekannt und führten zu direkten Kontakten.

Komplex Neue Erkenntnisse auf dem Gebiet der Hybrid-Mikroelektronik

Hier sind insbesondere die Vorträge zu erwähnen:

- Arbeiten zu hybridintegrierten Sensoren und optoelektronischen Hybrid-Anzeigemodulen auf der Grundlage einer stabilen Laborbasis (Ingenieurschule für Elektrotechnik und Keramik, Hermsdorf).
- Verfügbares Bauelemente-Sortiment des VEB Keramische Werke Hermsdorf: 1-Mbit-Hybridspeicher auf der Basis monolithischer 64-kbit- bzw. 256-kbit-Speicher.
- Neue Entwicklungsergebnisse zu Al_2O_3-Keramiksubstraten (UdSSR) und Übergabe von Mustern an Interessen zur Erprobung mit der Absicht, direkte bilaterale Kontakte herzustellen.

Vortragskomplex Technologie, Prüfung und Zuverlässigkeit elektronischer Bauelemente

Besondere Aufmerksamkeit fanden

- Ausführungen zum „ppm-Konzept" (DDR)
- Erläuterungen zum „Burn-in elektronischer Schaltkreise" (VRB)
- Vorträge zu modernen Verfahren für den Test und die Prüfung der Schaltkreisparameter, wie sie in der VRP und ČSSR gehandhabt werden, u.a. die Ausführungen zur automatischen Testeinrichtung TOA 1 für Schaltkreise (ČSSR)
- Diskussionen über spezifische BASIC-Varianten.

Neben diesen bauelementeorientierten Beratungskomplexen enthielt das Tagungsprogramm auch wieder solche anwenderbezogenen Arbeitsrichtungen der Mikroelektronik, wie

- Kleincomputer und Rechnerbaugruppen
- Rationalisierungs- und Automatisierungslösungen.

Erwartungsgemäß standen besonders die Ausführungen zu den Kleincomputern im Mittelpunkt des Interesses der Tagungsteilnehmer, z.B.:

- Programmiermodul EPROM 1
- NANOS-Gerätesystem
- Erläuterungen zum Einsatz und zu weiteren Zusatzbaugruppen des Kleincomputers KC 85
- Vorstellung des Bildungscomputers A 5105
- 16-bit-ESER-Personalcomputer 1834
- Industrie-Computer-Automatisierung (ICA).

Die Bildung von weiteren Nutzergemeinschaften für die typenspezifischen Gerätelinien wäre eine Möglichkeit zur effektiven Nutzung der Rechner als Rationalisierungsmittel. Die Vorführung der Geräte — ergänzend zu den Vorträgen — und praktische Erläuterung der Einsatzmöglichkeiten fanden einen großen Interessenkreis und führten zu zahlreichen vertiefenden Diskussionen. Dabei kann die Vorstellung des Bildungscomputers A 5105 als besonders gelungen angesehen werden.

Komplex „Automatisierungslösungen"

Es standen insbesondere Probleme der Einführung von CAD-, CAM- und CIM-Systemen im Mittelpunkt der Diskussionen. Großes Interesse fanden die Vorträge

- Rechnergestütztes Diagnostik- und Kontrollsystem für die MOS-LSI-Bauelementeproduktion „DIASTEMOS" (VRP; mit ergänzenden Teilsystemen auch im Rahmen der Posterdiskussion)
- Konzept zur rationellen Software-Entwicklung, das in der Ausbildung an der TU-Magdeburg genutzt wird.

Alle Vorträge zeichneten sich durch eine hohe Anschaulichkeit aus, wozu nicht zuletzt die Ausstellung moderner elektronischer Bauelemente aus dem Produktionsprogramm der Kombinate Mikroelektronik, Elektronische Bauelemente und Keramische Werke Hermsdorf (leider nicht das Sortiment vom VEB Carl Zeiss Jena) gehörte. In regen Fachdiskussionen spielte u.a. die Frage der Verfügbarkeit eine große Rolle. Die Nachfrage nach technischen Unterlagen konnte dankenswerterweise gut befriedigt werden. Ebenfalls positiv wirkte sich der Informationsstand zum staatlichen Informationsdienst „Mikroelektronik" aus, der vom VEB Applikationszentrum Elektronik Berlin eingerichtet wurde.

Schließlich ist nicht zuletzt die Software-Börse positiv zu werten, die vom VEB Datenverarbeitungszentrum Dresden betreut wurde. Der Speicher enthielt 1500 Software-Anwendungslösungen aus dem zentralen Speicher des DVZ, die mit AC 7150 und Drucker abrufbar waren. Weit über 100 Interessenten informierten sich vor allem für konstruktive CAD-/CAM-Lösungen und integrierte Systeme für 16-bit-Rechner.

Zusammenfassend läßt sich einschätzen, daß die 6. Fachtagung „Applikation Mikroelektronik" mit über 100 Vorträgen einen guten Überblick gab. Dabei wurde sowohl über den erreichten Stand berichtet als auch auf die nächsten Aufgaben orientiert, die zur Erhöhung der Leistungskraft der Volkswirtschaften mit Hilfe der Ingenieurorganisationen in den sozialistischen Ländern auf dem Gebiet der Mikroelektronik zu lösen sind.

NaK 53 *R. Schneider*

Berliner Informatik-Tage 1989

(Computers in Science and Higher Education)

a) 19. bis 23. Juni 1989
b) Berlin
c) Humboldt-Universität zu Berlin, Organisations- und Rechenzentrum
d) KI-Anwendungen, Problemlösungsumgebung und Wissenschaftlerarbeitsplätze, verteilte Verarbeitung, Leitung und Büroinformation, rechnergestützte wiss. Information und Publikation, Bibliotheksautomatisierung und wiss. Informationsdienste
Wissensch. Leitung: Prof. *Grabowski*
e) Humboldt-Universität zu Berlin
Organisations- und Rechenzentrum
Unter den Linden 6
Berlin, 1086

Information System, Work and Organization Design

Internationale wiss. Konferenz der IFIP

a) 10. bis 13. Juli 1989
b) Berlin
c) Humboldt-Universität zu Berlin, Sektion WID, IFIP-TC 9
d) Wechselwirkungen von Organisations-, Arbeits- und Informationssystemgestaltung — Software- und Orgwareentwicklung als Teil der Arbeits- und Organisationsgestaltung — Aufgaben sozialer Gruppen (Leiter, Spezialisten, ...) im Gestaltungsprozeß u.a.
Wissensch. Leitung:
Prof. *Fuchs-Kittowski*
e) Humboldt-Universität zu Berlin
Sektion WTO
August-Bebel-Platz, PSF 1297
Berlin, 1086

3. Tagung Multimikroprozessorsysteme „MMPS '89"

a) 6. bis 10. September 1989
b) Schwerin
c) Institut für Informatik und Rechentechnik der AdW der DDR
d) Architektur v. MMPS, Verbindungsstrukturen, Betriebs- und Programmiersysteme, Analyse und Simulation, Fehlertoleranz, Test und Debugging, Anwendungen in der Datenkommunikation und Prozeßsteuerung
Wissensch. Leitung: Prof. *Hammer*
e) Institut für Informatik und Rechentechnik der AdW der DDR
Rudower Chaussee 5
Berlin, 1199

III. Internationale wissenschaftliche Konferenz mit Ausstellung „Wissenschaftlicher Gerätebau und Automatisierung wissenschaftlicher Forschungsarbeiten" — Naucpribor RGW '89

a) 12. bis 15. September 1989,
vom 16. bis 20. September 1989 findet die öffentliche Geräteausstellung statt
b) Berlin, Kongreßhalle am Alexanderplatz
c) Akademie der Wissenschaften der DDR, Zentrum Wissenschaftlicher Gerätebau
d) Die III. Naucpribor folgt der Tradition der Vorgängerveranstaltungen in den Jahren 1978 in Moskau und 1984 in Plovdiv. Sie dient der Vertiefung der Zusammenarbeit zwischen Wissenschaftlern, Entwicklern, Konstrukteuren und Technologen für die Herstellung moderner Forschungstechnik in den RGW-Mitgliedsländern, insbesondere in den Akademien der Wissenschaften.
Die Zielstellungen sind:
   - die Demonstration der Zusammenarbeit auf allen strategisch wichtigen Gebieten bei der Entwicklung und Anwendung von Schlüsseltechnologien
   - die gegenseitige Information der RGW-Mitgliedsländer über ihre Leistungsfähigkeit durch die Darlegung wissenschaftlicher Forschungsergebnisse und die Demonstration von Erzeugnissen des wissenschaftlichen Gerätebaus und der Automatisierung wissenschaftlicher Forschungsarbeiten
   - die Ableitung von Schlußfolgerungen für die Erarbeitung nationaler, bi- und multilateraler Arbeitspläne für den Fünfjahrplanzeitraum 1991—1995
   - gegenseitige Bereitstellung von Forschungstechnik unter dem Gesichtspunkt des Ausbaus des RGW-Außenhandels und des Akademieaustausches sowie des Geräteaustausches zwischen den Hochschulen.

   Das wissenschaftliche Profil der III. Naucpribor richtet sich auf Geräte und Meßtechnik für Schlüsseltechnologien mit den Schwerpunkten:
   - Mikro- und Optoelektronik
   - Biotechnologie
   - Automatisierung wissenschaftlicher Forschungsarbeiten
   - Informationstechnik
   - Hochveredlung von Werkstoffen

Ausgehend von ihrer Zielstellung trägt die III. Naucpribor den Charakter einer wissenschaftlich-technischen Konferenz mit vorwiegend Plenarvorträgen. Eine Posterausstellung und die Gerätedemonstration ergänzen die Konferenz. Die Veröffentlichung der Vorträge ist in der Zeitschrift „Wissenschaftliche Apparaturen, Geräte und Automatisierungsmittel für die Forschung" vorgesehen.

Aufgabe der Veranstaltung ist es, die weitere Leistungsentwicklung auf dem Gebiet

des wissenschaftlichen Gerätebaus und der Automatisierung wissenschaftlicher Forschungsarbeiten sowohl innerstaatlich als auch im internationalen Maßstab entscheidend voranzubringen.
e) Teilnahme nur auf Einladung
Anfragen an:
Dr. sc. nat. *G. Peter* AdW der DDR, Zentrum Wissenschaftlicher Gerätebau, Rudower Chaussee 6, Berlin 1199

10. Fachtagung „Informationsverarbeitung bei konstruierender und projektierender Arbeitsweise"

a) 29. September 1989
b) Karl-Marx-Stadt
c) Technische Universität Karl-Marx-Stadt, Sektion Verarbeitungstechnik
d) • Konzipieren, Aufbauen und Anwenden moderner Informationsspeicher-, Informationsrecherche- und Informationsverarbeitungssysteme (Expertensysteme) im konstruktiven Entwicklungsprozeß
• Vermitteln methodischer Grundlagen zur Rationalisierung der Konstrukteurtätigkeit allgemein, speziell für kreative Tätigkeiten
• Einbeziehen moderner Methoden und Verfahren der geistig-schöpferischen Arbeit in die Lehre

Wissenschaftliche Leitung und Organisation: Dr.-Ing. *Preisler*, Tel.: 5612548
e) Technische Universität Karl-Marx-Stadt, Sektion Verarbeitungstechnik
PF 964
Karl-Marx-Stadt
9010
Teilnahmegebühr: 50,— M

International Workshop on Analogical and Inductive Inference (AII '89)

a) 2. bis 6. Oktober 1989
b) Reinhardsbrunn
c) Technische Hochschule Leipzig
Sektion Mathematik/Informatik
d) • induktive Software- und Hardwaresynthese
• mathematische und KI-Ansätze zum maschinellen (algorithmischen) Lernen, insbesondere analoges und induktives Schließen

Wissensch. Leitung: Prof. *Jantke*
e) Technische Hochschule Leipzig
Sektion Mathematik/Informatik
Eichendorfstr. 2
Leipzig, 7030

Aufsetzmontage in der Leiterplattenfertigung unter Berücksichtigung der Nacktchipverarbeitung

a) Die im Heft 5 (1988) für den 27. und 28. September 1989 angekündigte Fachtagung mit internationaler Beteiligung ist auf den 21. und 22. Dezember 1989 verschoben worden.
b) Berlin
c) Fachverband Elektrotechnik in der Kammer der Technik
Wissenschaftliche Sektion Technologie des elektronischen Gerätebaus
d) 1. Möglichkeiten und Grenzen der Baugruppenrealisierung
• monolithische Integration
• hybride Integration
• SMD-Technologie
• Nacktchipverarbeitung

2. Voraussetzung zur Einführung der Aufsetzmontage für SMD und Nacktchips
• Bauelementesortiment, Bauelementeverpackung
• Trägersysteme
• Hilfsmaterialien (Kleber, Lötpasten, Lote, Mikrodrähte, Verkappung)
• Bestückungstechnik
• Meß- und Prüftechnik

3. Verfahren und Ausrüstungen
• CAD von Leiterplatten u.a. Trägersysteme
• Fixierung und Kontaktierung der Bauelemente
• CAQ
• Reparaturmöglichkeiten der Baugruppen
• Spezialausrüstungen

4. Anwenderbeispiel
• SMD-Technik
• Nacktchipverarbeitung (einschl. Kundenwunschschaltkreise)

5. Fachspezifische Ausstellung
• Baugruppenmuster
• Sondergeräte
• Software für CAD, CAM, CAQ und weitere
• nachnutzbare Unterlagen.

e) Vortragsvorschläge und Ausstellungsangebote sind bis zum
31. Mai 1989
einzureichen;

Teilnahmemeldungen bitte schriftlich an folgende Anschrift:

Kammer der Technik
— Präsidium —
Fachverband Elektrotechnik
Clara-Zetkin-Str. 115/117
Berlin
1086

[1]) a) Termin, b) Veranstaltungsort, c) Veranstalter, d) inhaltliche Schwerpunkte, e) Anmeldungen an ...

Patente

Der Informationsbedarf im Forschungs- und Entwicklungsprozeß unter besonderer Berücksichtigung der Patentliteratur

Mitteilung aus dem Amt für Erfindungs- und Patentwesen der DDR

Im Wettlauf mit der Zeit kommt es mehr denn je darauf an, die verfügbaren Informationen über die in der Welt bereits bekannten technischen und technologischen Lösungen in allen Kombinaten und Forschungseinrichtungen konsequent in der Forschung und Entwicklung zu nutzen.

Um Ergebnisse zu erzielen, die über international bereits Bekanntes hinausgehen, ist es erforderlich, alle für die Bewertung des Standes der Technik in der Welt und die Entwicklung seiner Tendenzen vorhandenen Informationsquellen systematisch auszuwerten. Neben der Fachliteratur, den Standards, Prospekten, Messeberichten, Marktinformationen, Forschungs- und Entwicklungsberichten kommt der Patentliteratur dabei eine besondere Bedeutung zu. Die zwingende Notwendigkeit, den Intensivierungsfaktor „Information" umfassend zu nutzen, ist längst noch nicht in allen Forschungs- und Entwicklungskollektiven erkannt worden. Vielfach bestimmen historisch gewachsene Informationsgewohnheiten Qualität und Umfang der Literaturauswertung, weil Kenntnisse über den für die Lösung der Forschungs- und Entwicklungsaufgaben erforderlichen konkreten Informationsbedarf nicht vorliegen oder unzureichend sind sowie die erforderliche leitungsseitige Einflußnahme auf eine anforderungsgerechte Nutzung der Informationen über den Weltstand fehlt.

Durch das rasante Anwachsen des technischen Wissens, jährlich werden etwa

- 500 Tausend Fachbücher und Dissertationen
- 2 Mill. Artikel in 30000 Fachzeitschriften
- 300 Tausend Firmenschriften
- 100 Tausend Standards
- 100 Tausend Forschungs- und Entwicklungsberichte
- 1,2 Mill. Patentschriften

in der Welt veröffentlicht, wachsen kontinuierlich die Möglichkeiten, aber auch die Anforderungen an eine systematische Auswertung dieser Informationen [1].

Diese Informationsflut einerseits und die Maßstäbe, die an eine niveauvolle Arbeit der Wissenschaftler und Ingenieure andererseits zu stellen sind, erfordern Maßnahmen zu Veränderungen in der Bewertung des Rohstoffes „Information" und dazu, daß der tatsächliche Informationsbedarf zunehmend zum Ausgangspunkt der Arbeit von immer mehr Forschern wird [2].

▲ Die bei zahlreichen Forschern und Entwicklern noch bestehende Kluft zwischen den erforderlichen, den verfügbaren und den tatsächlich genutzten Informationen ist besonders groß bei der nichtbuchhändlerischen Literatur, insbesondere bei der Patentliteratur.

Ursache dafür ist eine gravierende Unterschätzung der Patentliteratur als wissenschaftlich-technische Informationsquelle. Dabei hat sie gegenüber anderen Informationsquellen deutlich sichtbare Vorteile:

• Patentinformationen sind die frühesten praxisbezogenen Veröffentlichungen über die in der Welt neuesten technischen und technologischen Lösungen, ihr Überschneidungsgrad mit anderen Informationsquellen ist gering (etwa 10%).

• Während Artikel in Fachzeitschriften oft allgemein und theoretisch sind, stellen Patentschriften immer konkrete technische Lösungen mit der Dreiteilung Darstellung des Standes der Technik, Kritik am Stand der Technik und Vorschlag zur Weiterentwicklung der Technik vor.

• In allen anderen Informationsquellen können durch geschickte Weglassung Mitbewerber und Konkurrenten getäuscht oder im Unklaren gelassen werden. Patenterteilungen setzen aber eine vollständige Offenbarung der neuen Lösungswege voraus.

• Die Patentliteratur offenbart in bezug auf die Vergangenheit das bereits bekannte

technische Wissen, bezüglich der Gegenwart die neuesten technischen und technologischen Lösungen und in bezug auf die Zukunft die absehbaren Tendenzen im Wandel des wissenschaftlich-technischen Fortschritts.

Die damit nur kurz angedeutete besondere Bedeutung der Patentliteratur als wissenschaftlich-technische Informationsquelle erfordert, das in Patenten gespeicherte Wissen konsequent für ein hohes Niveau der Arbeit in Forschung und Entwicklung auszuwerten. Dabei gilt es in allen Phasen der Forschungs- und Entwicklungsarbeit, von der strategischen Arbeit bis zur Vorbereitung der Markteinführung, diesen riesigen Wissensschatz – der Welt-Patentfonds umfaßt gegenwärtig etwa 35 Mill. Dokumente – als Ideenreservoir und unbestechlichen Gradmesser für die eigenen Forschungsergebnisse zu nutzen.

Informationsbedarf für die strategische Arbeit

Die Erarbeitung strategischer Zielstellungen erfordert, aus den sich international abzeichnenden Tendenzen die bestimmenden Entwicklungstendenzen abzuleiten und Aufgaben festzulegen, die über international bereits Bekanntes hinausgehen. Da die Patentliteratur über die weltweit fortgeschrittensten technischen und technologischen Lösungen, über strukturelle Veränderungen in Forschung und Entwicklung sowie über die Marktinteressen der führenden Konzerne und Einrichtungen informiert, sind zur Deckung des Informationsbedarfs folgende Patentauswertungen erforderlich:

- statistische Untersuchung der Dynamik der sich weltweit vollziehenden Entwicklungen auf dem Fachgebiet
- statistische Analyse der länderspezifischen Entwicklung des internationalen Patentgeschehens
- Ermittlung und Bewertung des Patentbesitzes von Marktführern und Konkurrenten
- Recherchen zum Stand der Technik in den Staaten, die den Weltstand bestimmen sowie in den vorgesehenen Exportländern
- Bewertung von Schlüsselpatenten.

Die strategische Arbeit wird immer mehr bestimmt vom hohen Tempo des wissenschaftlich-technischen Fortschritts und von der zunehmenden Verflechtung technischer Gebiete.

Informationsgrundlagen dafür sind die internationale Patentstatistik, die Patentfonds in den Kombinaten und Forschungs- und Entwicklungseinrichtungen, den Polytechnischen Patentbibliotheken sowie der Zentralen Patentbibliothek des AfEP.

Informationsbedarf für die Ausarbeitung und Realisierung anspruchsvoller Zielstellungen von Pflichtenheften

Die Ausarbeitung anspruchsvoller Zielstellungen von Pflichtenheften ist in hohem Maße abhängig von einer genauen Kenntnis der volkswirtschaftlichen Anforderungen, der Marktbedingungen, der Bedarfsentwicklung und des Weltstands. Aus der Patentliteratur lassen sich am sichersten Zielstellungen für Pflichtenhefte zu international neuen technischen und technologischen Lösungen ableiten, wodurch die ökonomischen Effekte entscheidend beeinflußt werden.

Zur Vorbereitung des Pflichtenheftes haben sich folgende Patentauswertungen in den Fonds der wichtigsten Industriestaaten bewährt:

- Recherchen zum Stand der Technik in den führenden Industriestaaten zur Ermittlung objektbezogener Informationen zum Weltstand
- konzernbezogene Analysen zur Feststellung der technischen Entwicklungsrichtungen und des Patentbesitzes von Marktführern und Konkurrenten zur Ableitung der eigenen Forschungsziele, die auf international neue Weise zu lösen sind
- Bewertung der volkswirtschaftlich besonders bedeutsamen Erfindungen des Fachgebiets.

In allen Entwicklungsphasen sind ergebnisorientierte Vergleiche der erreichten Forschungsleistungen mit dem in der Patentliteratur dokumentierten Weltstand erforderlich.

Folgende weitere Patentauswertungen sind durchzuführen, wenn im Vorfeld einer Patentanmeldung im AfEP die Schutzfähigkeit der erreichten Forschungsergebnisse im Kombinat zu bewerten ist:

- Recherche im internationalen Patentfonds zur Ermittlung von Erfindungsbeschreibungen, in denen die Merkmale der eigenen Lösung vollständig oder teilweise identisch oder äquivalent beschrieben sind
- merkmalorientierter Vergleich des wissenschaftlich-technischen Gehalts der Schutzwirkungen der ermittelten relevanten Erfindungsschutzrechte mit den eigenen wissenschaftlich-technischen Ergebnissen.

Diese Patentrecherchen müssen in den Referate- und Volltextfonds der Kombinate und Forschungs- und Entwicklungseinrichtungen, der Polytechnischen Patentbibliotheken oder der Zentralen Patentbibliothek des AfEP durchgeführt werden.

Themenbezogene Überwachungen technischer Gebiete vom Zeitpunkt der Eröffnungsverteidigung bis zur Einführung in die Produktion werden durch die mit dem Welt-Patent-Dienst bereitgestellten Neuveröffentlichungen zu technischen Profilen und zu Anmeldern (Inhabern und Erfindern) sowie Informationen des AfEP, die aus der Überwachung der Entwicklung des territorialen Schutzumfangs ausgewählter Erfindungen und der Änderungen des Rechtsbestands wichtiger Schutzrechte resultieren, unterstützt. Die Darstellung der in den verschiedenen Phasen des Forschungs- und Entwicklungsprozesses erforderlichen Patentauswertungen zeigt, daß der Informationsbedarf aus Patenten eine sich im Verlauf der Forschungsarbeit dynamisch verändernde Größe ist. Der objektiv notwendige Informationsbedarf, der sich im Bewußtsein der Forscher- und Entwickler subjektiv widerspiegelt, kann keine feste Größe sein, weil ständig Informationen verarbeitet werden, neue Informationen entstehen und dadurch wieder ein neuer Informationsbedarf entsteht.

Das erfordert, Recherchestrategien festzulegen, die eine gründliche Information über das in der Patentliteratur gespeicherte Wissen zu den in der Welt bereits bekannten technischen und technologischen Lösungen sichern.

Dabei muß sich die Deckung des Informationsbedarfs aus Patenten zunehmend durch die Integration der Patentinformation in den unmittelbaren Forschungs- und Entwicklungsprozeß vollziehen. Dazu wird die Voraussetzung, die Nutzung der Rechentechnik für die Patentinformation, insbesondere der schnelle Fernzugriff zu den Patentdatenbanken des AfEP und die erweiterten Auswertemöglichkeiten in der Nähe der Arbeitsplätze der Nutzer schrittweise geschaffen [3] [4].

Für die Deckung des Informationsbedarfs stehen weiterhin die Leistungen der Patentinformationsdienste in den Kombinaten und Patentbibliotheken der Bezirke Dresden, Karl-Marx-Stadt, Magdeburg, Rostock und Suhl, der Zentralen Patentbibliothek in Berlin sowie die konventionellen und rechnergestützten Dienste des AfEP zur Verfügung.

NaA 41 *J. Zimdars*

Literatur

[1] *Cuhausz, H. B.:* Online-Recherchen in externen Technik Datenbanken. Technische Mitteilungen 81 (1988) 1, S. 3

[2] *Rückl, St.; Schmoll, G.:* Lexikon der Information und Dokumentation. Leipzig: Bibliographisches Institut 1984, S. 188

[3] *Janke, L.:* Patentdatenbanken des AfEP. der neuerer 37 (1988) 5, S. 66–68

[4] *Zimdars, J.:* Rechnereinsatz für die Patentinformation. der neuerer 37 (1988) 6, S. 86–87

Regelschaltung für den Austastwert in Bildsignalen

DD-WP:	262 314
Int.-Cl.:	H 04 N 5/16
Anmeldetag:	13. 07. 87
Anmelder:	Deutsche Post, Berlin

Die Erfindung bezieht sich auf das Gebiet der Farbfernsehstudiotechnik und betrifft die Analog-Digital-Umsetzung von Farbbildsignalen in Dreikomponentendarstellung bzw. von vollständigen Farbbildsignalen. Sie ist im Digitalcoder von digitalen Studiogeräten, z. B. Standbildspeichern, anwendbar.

Bei der AD-Umsetzung von Farbbildsignalen bilden die OIRT-Empfehlung Nr. 106 und die CCIR-Empfehlung Nr. 601 u. a. die Grundlage für die Zuordnung der Quantisierungspegel zu den Bildsignalanteilen. Dabei sind die Austastwerte (AW) des Leuchtdichtesignals Y und der Farbdifferenzsignale $B-Y$ und $R-Y$ relevante Bezugspegel. Dem Austastwert für Y (AWY) ist die 16 und den Austastwerten für die digitalen Farbdifferenzsignale C_B und C_R (AWC) ist die 128 zugeordnet. Damit die analogen Austastwerte am analogen Eingang des ADU über die notwendige Pegelstabilität verfügen, müssen im analogen Teil schaltungstechnische Maßnahmen den analogen Austastwert möglichst konstant halten.

Eine Analyse der technischen Mängelursachen zeigt, daß die ausschließlich auf Stabilisierung des analogen Austastwerts ausgerichteten Lösungen nicht den hochstabilen Austastwert DAW gewährleisten, der angestrebt ist. Deshalb besteht die Aufgabe der Erfindung darin, eine andere geeignetere Möglichkeit für die Regelung des Austastwertes unter Einbeziehung des Analog-Digital-Umsetzers zu finden.

Diese Aufgabe löst die Erfindung auf über-

raschend einfache Weise dadurch, daß die Regelschleife über den Analog-Digital-Umsetzer hinweg zurückgeführt ist und eine mit dessen Eingangstakt und Zeilenfrequenz getastete Vergleichsschaltung für das *n* bit breite digitale Bildsignal und einen vorprogrammierten digitalen Austastwert sowie eine digital-analog-umsetzende Schaltung enthält. In vorteilhafter Weiterbildung der Erfindung können weitere digitale und/oder analoge Schaltkreise mit erfaßt sein.
Der Effekt der Erfindung besteht darin, daß durch die Einbeziehung des ADU in den Regelkreis und die Eingabe digitaler Werte für den jeweiligen Austastwert eine hochstabile Regelung dieses relevanten Bezugspegels möglich ist.

NaP 135

Adreßgenerator zur Fehlerortung auf PCM-Trakten

DD-WP: 262 524
Int.-Cl.: H 04 B 17/02
Anmeldetag: 02. 06. 86
Anmelder: VEB Nachrichtenelektronik Greifswald

Das Ziel der Erfindung ist ein Generator zur sicheren Adressierung der Zwischenstellen und der erforderlichen Schleifenbildung für die Fehlerortung defekter Regeneratoren in PCM-Leitungstrakten, der trotz geringem technischem Aufwand und unkomplizierter Bedienbarkeit eine hohe Zuverlässigkeit garantiert.
Aufgabe der Erfindung ist es deshalb, einen Adreßgenerator zu schaffen, der die für die Ortung defekter Regeneratoren im Leitungstrakt von PCM-Übertragungseinrichtungen notwendigen Zyklen zur Steuerung des Prüfschleifenschalters automatisiert und den erfolgten Schleifenschluß quittiert.
Erfindungsgemäß wird diese Aufgabe wie folgt gelöst: Der Adreßgenerator zur Fehlerortung defekter Regeneratoren in adressierbaren Zwischenstellen von PCM-Leitungstrakten liefert mit Hilfe von Schleifenverfahren automatisch den Zyklus der Zustände während des Ortungsverfahrens in der Reihenfolge

- Löschen aller geschlossenen Schleifen
- Vorbereiten als Schutz gegen ungewolltes, zufälliges Erkennen einer Adresse einer Schleife
- Schleifenschluß der adressierten Zwischenstelle und unterschiedlich langen Impulssignals

in Form eines Steuersignals aus 4 zeitlich hintereinander folgenden, aber unterschiedlich langen Impulsgruppen, die durch konstante Pausen getrennt sind, an den Leitungssignalcoder der PCM-Anlage. Dieses erfolgt im Adreßgenerator mit einer Zyklussteuerungsschaltung, die aus einem Johnson-Zähler mit nachgeordneten logischen Gattern zur Auswertung des Zählerzustands besteht. Der Ausgang dieser Zyklussteuerungsschaltung ist einerseits mit dem Steuereingang eines Datenmultiplexers, an dem der Adressenwahlschalter angeschlossen ist, und mit den Dateneingängen des Datenmultiplexers verbunden. Dabei ist die Anzahl der Dateneingänge eine Funktion der erforderlichen Adressenzahl für die ortenden Zwischenstellen. Die Datenausgänge des Datenmultiplexers sind an den Rückwärtszähler mit nachgeschalteter Auswertelogik geführt. Der Ausgang der Auswertelogik ist wiederum an den Ladeeingang des Rückwärtszählers als auch an den Zähleingang der Zyklussteuerungsschaltung angeschlossen. Ein weiterer Ausgang der Zyklussteuerungsschaltung ist mit dem 1. Eingang einer Generatorsteuerlogik, deren 2. Eingang mit dem 1. Ausgang einer Zyklusanlaufschaltung, und deren Ausgang mit dem Steuereingang eines *RC*-Generators verbunden sind.
Der 2. Ausgang der Zyklusanlaufschaltung ist über die Torschaltung geführt, die mit der Endstufe des erfindungsgemäßen Adreßgenerators verkoppelt ist. Das vom Adreßgenerator und über die aufgebaute Schleife in der angewählten Zwischenstelle zur ortenden PCM-Anlage zurückgeleitete Quittierungssignal wird von einer Quittierungsschaltung im Adreßgenerator registriert.

NaP 140 *D. Ufert*

Buchbesprechungen

Regelungstechnik auf dem Personalcomputer
Von *M. Peschel*
Berlin: VEB Verlag Technik, 1. Auflage 1988
206 Seiten, 102 Bilder, 40 Programme, Lederin, 22,— M

Die moderne Mikrorechentechnik bestimmt in starkem Maße den wissenschaftlich-technischen Fortschritt. Daher ist es erklärlich, daß viele Bücher über Arbeitsweise und Programmierung von Rechnern erschienen sind und guten Absatz haben. Experimentierbücher zum Training des wissenschaftlichen Arbeitens auf dem Mikrorechner und zum spielenden Erkenntnisgewinn gibt es hingegen nur wenige. Vorrecht eines solchen Buches ist das Eingehen auf interessante Fragestellungen, ohne die Vollständigkeit eines Lehrbuches zu erreichen.
In dem Buch werden zunächst allgemeine Fragen zu den Grundlagen der Regelungstechnik behandelt: Beziehungen zur Philosophie, das Wechselwirkungs- und das Rückkopplungsprinzip mit den möglichen Folgen für das Verhalten der Systeme.
Systemtheoretische Grundlagen zur Betrachtung des Verhaltens von Gliedern werden in Theorie und Computer-Experiment betrachtet.
Den linearen Regelungssystemen und Netzwerken gilt das nächste Kapitel, wobei Eigenarten bei verschiedenen eingesetzten Reglertypen mit dem Computer-Experiment erforscht werden. Bei nichtlinearen Regelungssystemen interessieren besonders moderne Konzepte, wie z.B. das vom Verfasser erarbeitete Evolon-Konzept oder adaptive Mikrorechnerregler.
Einige Anwendungen beschließen das Buch, wobei hier besonders neue Erkenntnisse vorgestellt werden, die manchmal den Charakter von Originalarbeiten haben, wie z.B. die Theorie ratengekoppelter Regelungssysteme, Chaos in Regelkreisen, Hyperzyklen.
Es werden 89 Programme in SIMONS-BASIC zum Experimentieren vorgegeben, wobei bei jedem Programm einige typische Zeitverläufe zur Kontrolle, ob das Programm richtig „läuft", angegeben werden. Die Programme können leicht in das BASIC anderer Mikrorechner umgeformt werden.
Das Buch wendet sich in erster Linie an Anfänger, z.B. Studenten, interessierte Oberschüler (die Programme wurden tatsächlich bei einer Interessengemeinschaft an einer Berliner EOS ausprobiert) aber auch an Fachleute, die sich spielend in das nicht einfache Gebiet der modernen Regelungstechnik einarbeiten wollen. Das Buch dient dem Programmier-Training, gibt aber auch eine große Anzahl von Anregungen für die interessante Aufgabe der Gewinnung völlig neuer Erkenntnisse. Insofern ist es mehr als ein Leitfaden für Anfänger.

NaB 831 *K. Göldner*

Angewandte Akustik Bd. 1
Von *W. Kraak* und *G. Schommartz*
Berlin: VEB Verlag Technik, 1. Auflage 1987
284 Seiten, 149 Bilder, 10 Tafeln, Broschur, 35,— M

Mit dem vorliegenden ersten Band wird die neue Buchreihe „Angewandte Akustik" eröffnet, in der Autoren aus allen Teilgebieten dieser Fachdisziplin über ein aktuelles Forschungsgebiet berichten oder einen Überblick über einen größeren Forschungskomplex geben.
Aufgrund des hochgradig interdisziplinären Charakters der Akustik stehen physikalische, maschinentechnische, architektonische, medizinische, elektrotechnische, musikalische oder signalanalytische Aspekte im Vordergrund oder werden in unterschiedlichen Verknüpfungen zueinander dargestellt. Den renommierten Herausgebern *W. Kraak* und *G. Schommartz* sowie den Autoren der einzelnen Beiträge ist es bereits im 1. Band dieses Jahrbuches gelungen, einen Eindruck von der Fülle des Stofflichen und Methodischen zu geben, das derzeit im Vordergrund akustischer Grundlagenforschung und ihrer Applikation steht. Der thematische Bogen spannt sich von Arbeiten auf klassischen Gebieten (z.B. *W. Fasold*, *H. Winkler* „Realisierung raumakustischer Forderungen in neuen Saalbauten") über Beiträge zu modernen Entwicklungen der Meßtechnik (*J. Plundrich* „Schallintensitätsmeßtechnik") und Signalanalyse (*V. Naumburger* „LPC — Eine Signaltransformation für die Digitaltechnik") bis zu Abhandlungen über Einsatzgebiete der Akustik in der Mikroelektronik, Kerntechnik und Geophysik (z.B. *R. Apel* „Akustische Verfahren in der Geophysik").
Es ist besonders hervorzuheben, daß das vorliegende Werk nicht nur einen kleinen Kreis von Spezialisten anspricht, sondern sicher aufgrund des zum großen Teil allgemeinverständlich gehaltenen „roten Fadens" in jedem Artikel eine Vielzahl technisch interessierter Laien erreichen wird. Darüber hinaus trägt es nach Meinung des Rezensenten wirkungsvoll dazu bei, daß — wie im Geleitwort vermerkt — die Nachbarn (d.h. die Akustiker der verschiedenen Fachbereiche) über ihre Zäune schauen.
Zusammenfassend kann festgestellt werden, daß die neue Buchreihe, die selbstverständlich kein Ersatz für die immer noch in der DDR fehlende nationale Akustikzeitschrift sein kann und auch nicht will, einen guten, auch in ansprechender Form präsen-

tierten Überblick über die Aktivitäten auf dem Gebiet der Akustik in unserem Land gibt. Einer wünschenswerten internationalen Verbreitung dieses Jahrbuchs steht jedoch die durchgängig deutschsprachige Abfassung der Texte entgegen, die auch nicht durch vorangestellte englische und französische Kurzfassungen kompensiert werden kann. Im Sinne einer Erhöhung der Reputation der Akustik der DDR in den führenden Industrieländern sollte dieser gravierende Nachteil bei den folgenden Bänden überwunden werden.

NaB 830 *D. Hamann*

Grundgebiete der Elektrotechnik Band 1 Stationäre Vorgänge

Von *A. Führer, K. Heidemann* und *W. Nerveter*
München, Wien: Carl Hanser Verlag, 3., verbesserte Auflage 1988, 268 Seiten, 484 Bilder, Broschur

Die beiden Bände über die Grundlagen der Elektrotechnik sind vorrangig als Lehrbücher für Studierende und als Nachschlagewerk und Leitfaden für Techniker und Ingenieure gedacht. Im vorliegenden Band 1 werden stationäre Vorgänge in elektrischen Stromkreisen und in elektrischen und magnetischen Feldern behandelt.
Das Buch gehört zur Reihe „Studienbücher der technischen Wissenschaften".
Bei den Lesern bzw. Studierenden werden elementare Kenntnisse in der Algebra, der Differential- und Integralrechnung sowie der Trigonometrie vorausgesetzt. Grundlagenlehrbücher unterscheiden sich, einen einwandfreien fachlichen Gehalt vorausgesetzt — was in diesem Fall gegeben ist, vor allem in der didaktischen Art, in der der Lernende an den Stoff herangeführt wird, im logischen Aufbau, im exakten Gebrauch der Begriffe und in der Verständlichkeit der Darstellung.
Einleitend wird dem Lernenden ein Vorschlag unterbreitet, sich den Stoff in 3 Stufen zu erarbeiten:

Kennenlernen — Durcharbeiten — Vertiefen. Mit diesem Vorschlag verfügt der Leser über eine hervorragende Arbeitsregel, die nicht nur empfehlenswert ist, sondern die durch den inhaltlichen Aufbau und die Art der Darstellung in hervorragender Weise unterstützt wird. Die Hauptkapitel sind:

1. Grundbegriffe
2. Zweipole
3. Vierpole
4. Netzwerkanalyse
5. Das elektrische Feld
6. Das magnetische Feld
7. Leitungsmechanismen.

Exaktheit der Begriffe und Verständlichkeit der Darstellung sind erfreuliche Kennzeichen dieses Lehrbuchs. Dazu gehört auch, daß der Begriff „Bezugspfeil" besonders herausgearbeitet wurde und das „Verbraucherpfeilsystem" bei der Darstellung angewendet wird.
Jedem Unterkapitel sind zunächst die Lernziele vorangestellt. Es folgen jeweils: die reich bebilderte Darstellung des Inhalts mit hervorgehobenen Definitionen und Kernaussagen; die leicht verständlichen Beispiele; der „Praxisbezug"; die das eigene Nachdenken anregenden Fragen und die mit Schwierigkeitsgrad (3 Stufen) gekennzeichneten Aufgaben (Lösungen im Anhang!).
Erfreulich ist die hervorragende, das Lernen und die Übersichtlichkeit unterstützende Gestaltung, die nahezu alle Möglichkeiten nutzt, ohne zu überfrachten.
Konsequente Beschränkung auf das Wesentliche, Exaktheit, Verständlichkeit und die geschickte unauffällige Didaktik sind nicht nur für den Einsteiger in die Elektrotechnik sehr hilfsreich, sondern das Buch ist auch für Ingenieure in der Praxis hervorragend geeignet, das Wissen aufzufrischen und im Bedarfsfall nachzuschlagen. Die Angabe der wichtigsten Fachausdrücke in englischer Sprache hilft den angehenden Technikern beim Erlernen der englischen Fachsprache.

NaB 829 *E.-O. Schönemann*

Grundlagen der Halbleiter-Elektronik

Von *R. Müller*
Reihe „Halbleiter-Elektronik" Bd. 1
Berlin: Springer-Verlag, 5. Auflage 1987
203 Seiten, 123 Bilder, Broschur

Band 1 der insgesamt etwa 20 Bände umfassenden Reihe soll laut Vorwort zur 1. Auflage von 1971 „den Leser auf möglichst einfache Weise in die Lage versetzen, Fachliteratur über Halbleiterbauelemente mit der nötigen Kritik lesen zu können".
Einleitend werden hierzu inzwischen der Schulphysik zuzuordnende Halbleitergrundbegriffe, wie Eigen- oder Störstellenleitung sowie Ladungsträgergeneration oder -rekombination, anhand einfacher Aufbau- und Bindungsvorstellungen für ein Einzelatom im isotropen Festkörper eingeführt. Das prinzipielle Bändermodell eines Halbleiters wird im weiteren durch Berücksichtigung der konkreten Wechselwirkung benachbarter Gitterbausteine im realen Kristall verfeinert. Unter Zuhilfenahme des quantenmechanischen Wellenzahlvektors entstehen für ausgewählte Kristallrichtungen die Bänderstrukturen der wichtigsten Halbleitermaterialien, die ihre Klassifizierung als direkte oder indirekte Halbleiter ermöglichen.
In den folgenden Kapiteln werden Störungen des thermodynamischen Gleichgewichts der Ladungsträgerkonzentrationen betrachtet, deren dynamischen Abbau die Zeitkonstanten „dielektrische Relaxation" bzw. „Minoritätsträgerlebensdauer" bestimmen. Die Quasi-FERMI-Niveaus (Imref) für Elektronen sowie für Löcher werden eingeführt.
Abschließend werden für den pn-Übergang („das Grundelement der meisten Halbleiterbauelemente") insbesondere auch Sperrschicht- und Diffusionskapazität betrachtet sowie Varianten des Sperrspannungsdurchbruchs.
Geschickt eingeflochten werden — jeweils am Ende der 7 Kapitel — verbal zu beantwortende Fragen oder Kontrollaufgaben, die dem Lernenden immer wieder aufzeigen, was er gerade überlesen hat.
Rückgriffe auf Physik- oder Mathematikbücher werden dem Leser zum Teil durch den Anhang erspart — z.B. mit 7 bzw. 8 Seiten über Teilchen und Wellen oder Verteilungsfunktionen.
Der Autor hat erfreulicherweise keine Scheu vor sehr anschaulichen Vergleichen. So läßt er z.B. Löcher als „Luftblasen" im „Flüssigkeits"-Valenzband zu dessen Oberkante aufsteigen.
An der 5. Auflage eines eingeführten Buches gibt es naturgemäß wenig zu kritisieren. Zu diesem Wenigen aber zählt sicher, daß in der vorangestellten Auflistung z.B. der SI-Basiseinheit Ampere die Dimension $T^{-1} \cdot Q$ zugeordnet wird, d.h. quasi ein MKSC-Vierersystem propagiert wird. Ersparen wir dem angehenden Ingenieur die überwundenen Maß- bzw. Einheitensysteme des Physikers!
Das Studium des Buches gibt dem Leser den angestrebten Einblick in die physikalischen Grundlagen der Halbleiterelektronik. Damit wird das Verständnis der Vielzahl ihrer Bauelemente sowie der technischen und technologischen Probleme erleichtert.

NaB 821 *W. Michel*

Kurzinformationen

• Mit einer Frequenz von 22 GHz läßt sich das Licht einer **Indium-Gallium-Arsenid-Laserdiode** modulieren, die von Forschern der GTE Laboratories in Waltham (Massachusetts, USA) entwickelt wurde. Der bisherige Rekord stand bei 18 GHz. Die Steigerung wurde erreicht, indem die Dicke der lichtemittierenden Schicht von 1,5 µm auf 0,8 µm verringert wurde. Die Bandbreite würde ausreichen, um über eine LWL-Verbindung 200 bis 400 Videokanäle zu übertragen — vorausgesetzt, es finden sich Glasfasern und Empfänger, die hier noch mitspielen. Für einen praktischen Einsatz sind noch einige Jahre weiterer Entwicklungsarbeit notwendig. Zunächst arbeitet die Firma an einem 12-GHz-System für 120 Videokanäle.

• Der Computerkonzern IBM arbeitet an einem als Röntgen-Lithographie bezeichneten Verfahren, mit dem die **Herstellung von 64-Mbit-Speicherchips** möglich werden soll. Das wäre das 16fache der Dichte, die zur Zeit bei den leistungsfähigsten, noch nicht in Massenproduktion hergestellten Chips erreicht wird. Mit der neuen Methode könnten künftig mehrere übereinander liegende Schichten ihre Schaltkreisgravur erhalten, ohne daß der Strahl wie bei anderen Verfahren immer wieder über den Chip geführt werden müsse, um Linie für Linie zu erzeugen. Ein weiterer Vorteil bestehe darin, daß Röntgenstrahlen Staubpartikel auf den Masken durchdringen, die bei der Nutzung von Strahlen größerer Wellenlänge zu Ausschuß führen. Probleme bei der Röntgen-Methode seien die schwierige Konzentration der Strahlen im gewünschten Brennpunkt und die weitere Miniaturisierung von Trägern und Masken. (ADN)

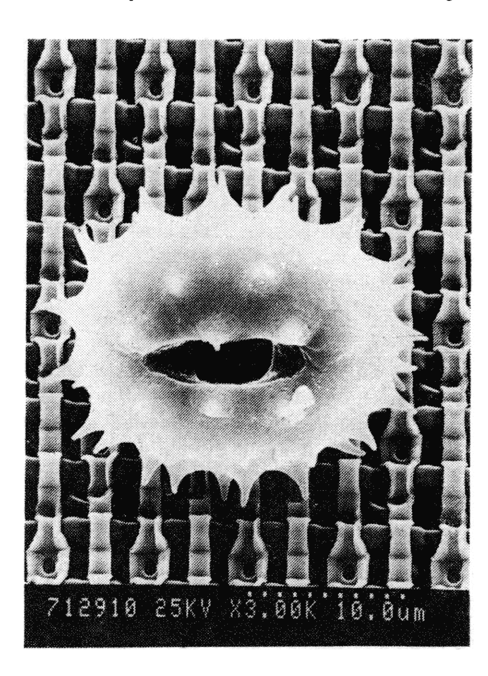
• **Teilnehmernetz zur Datenübertragung über Satelliten:** Die Datenübertragung zwischen Teilnehmern eines regionalen Netzes über Satelliten wird zur Zeit international

diskutiert, da sie Vorteile gegenüber anderen Übertragungsarten hat. In den USA wurde ein derartiges System geschaffen und in Westeuropa laufen Untersuchungen an einem entsprechenden Netzwerk. Wesentlich bei einem solchen System ist, daß bereits vorhandene Satelliten benutzt werden und die funktechnischen Einrichtungen der Terminals bei den einzelnen Teilnehmern durch die Verwendung einer Antenne mit kleiner Fläche (Apertur) geringe Abmessungen haben und die gesamte Anlage billig ist. Daher werden diese Terminals mit VAST (very small aperture terminals) bezeichnet. Es gibt 2 verschiedene Übertragungsarten:

- interaktiver Datenverkehr; jeder Teilnehmer kann Daten über den Satelliten senden und empfangen
- Datenverteilung; von einer zentralen Stelle (Hauptstation) werden die Daten über den Satelliten gesendet und von den Teilnehmern (Nebenstationen) empfangen.

Das in den USA betriebene System dient der Datenverteilung im Finanzwesen. Die Untersuchungen in Westeuropa ordnen einem interaktiven Datenverkehr die größere Bedeutung zu. Dabei ist vorgesehen, für die Übertragung von der Hauptstation zum Satelliten eine Bitrate von 19,6 kbit/s anzusetzen. Durch eine Adressierung können Teilnehmer selektiv verbunden werden. Die Anzahl der Teilnehmer einer Teilnehmergruppe (regional begrenzt) ist durch die Aufteilungsquote der 19,6 kbit/s gegeben. Als geeignete Satelliten kommen in Europa INTELSAT V, EUTELSAT I und TV-SAT II in Frage. Die erforderlichen Antennendurchmesser für die Teilnehmerterminals liegen bei 60, bzw. 30, bzw. 10 cm. Die Kosten eines Terminals werden etwa 7500 bis 25000 $ betragen. Für die nächsten 10 Jahre wird geschätzt, daß etwa 500000 VSAT-Anlagen weltweit installiert werden. *W. Mansfeld*

• Die Chipfläche seines **1-Mbit-DRAM** hat Siemens jetzt von 54 auf 45 mm² „geshrinkt". Das Foto (Siemens-Pressefoto) zeigt einen 0,001 mm² „großen" Chipausschnitt mit der Blütenpolle eines Gänseblümchens. Deutlich sind die Leiterbahnen mit kleinsten Strukturen von 1,0 µm zu sehen. Diese Speicherbausteine haben exakt 1048576 bit. Von der Mega-Technologie sollen Produkte wie ISDN-Bausteine oder Chips für digitale TV-Geräte sukzessiv profitieren. Wachsende Chipflächen werden bis Mitte der 90er Jahre 500 mm² erreichen.
Für 1989 plant Siemens, die Fertigungsstückzahl seines 1-Mbit-DRAM (HYB 511000) deutlich zu steigern (1988: 1 Mio Stück). Dazu wird die vorhandene Fertigungsfläche (je 4000 m² für Reinstraum bzw. Montage/Prüffeld) vollständig ausgenutzt. Auch das „Equipment" wird erweitert.



• **Bitfehler- und Jittermeßplatz in einem Gerät:** Die kompakten Bitfehler- und Jittermeßgeräte PF-8/PFJ-8 für den universellen Einsatz bei der Entwicklung, Fertigung, Installation und Fehlersuche an Übertragungssystemen stellte Wandel & Goltermann auf der „electronica" erstmals vor.
Diese Geräte sind besonders für die Messung von Fehlern an Signalen mit Rahmenstruktur geeignet, wie sie bei digitalen Übertragungsstrecken vorliegen. Im Bereich von 48 kbit/s bis 11 Mbit/s messen die Geräte Bit-, Block-, Code- und Rahmenfehler. Der Meßplatz PFJ-8 erzeugt und mißt zusätzlich noch Jitter und nimmt die Jitterverträglichkeit und Jitterübertragungsfunktion eines Prüflings, zum Beispiel eines Demultiplexers, automatisch auf.
Der Sender der Geräte kann einstellbare Fehlerhäufigkeiten in Meßsignale einblenden, der des PFJ-8 zusätzlich die Sendesignale mit definiertem Jitter versehen. Ausgewertet werden die gemessenen Fehler u.a. nach der internationalen CCITT-Empfehlung G.821.
Weitere Eigenschaften, die dazu beitragen, die Geräte zu universell einsetzbaren Fehlermeßgeräten zu machen, sind die Messung der Signallaufzeit und der sehr geringe Eigenjitter.
Beide Geräte verfügen über einen eingebauten Drucker und sind über eine IEC-/IEEE-Bus-Schnittstelle fernsteuerbar.

• Die Internationale Elektrotechnische Kommission (IEC) wird bald ein neues Kapitel des **„Internationalen elektrotechnischen Wörterbuches"** herausgeben, das Begriffe und Definitionen der Telekommunikationstechnik enthalten wird. Das Kapitel 701, 57 Seiten im A4-Format, wird aus 3 Teilen bestehen:

- Telekommunikationsordnungen
- Schaltungen und Netzwerke
- Nutzung und Betrieb von Schaltungen und Netzen.

Aufgelistet sind Begriffe und Definitionen in Englisch, Französisch, Russisch und Spanisch sowie die Begriffe in Deutsch, Italienisch, Holländisch, Polnisch und Schwedisch ohne Definitionen. Ein Index in allen 9 Sprachen ist angefügt.
Andere bereits verfügbare Kapitel aus der gleichen Serie sind:

- Kapitel 725 Satellitenfunktechnik
- Kapitel 726 Übertragungsleitungen und Wellenleiter.

Nachrichtentechnik · Elektronik

Herausgeber: Kammer der Technik

Verlag: VEB Verlag Technik
Oranienburger Str. 13–14, Berlin, DDR, 1020
Telegrammadresse: Technikverlag Berlin
Telefon: 28700: Telex: 0112228 techn dd

Verlagsdirektor: Dipl.-Ing. *Klaus Hieronimus*

Redaktion: Dipl.-Ing. *Ernst-O. Schönemann*, Verantwortlicher Redakteur (Telefon: 2870332, 2870327) Ing. *Gisela Rehahn*, Redakteur (Telefon: 2870347)

Gestalter: *Irmgard Greschek* (Telefon: 2870339)

Lizenz-Nr.: 1103 des Presseamtes beim Vorsitzenden des Ministerrates der Deutschen Demokratischen Republik

AN (EDV): 18432

Gesamtherstellung: Messedruck Leipzig, BT Borsdorf, Borsdorf, 7122, III-18-328

Anzeigenannahme: Für Bevölkerungsanzeigen alle Anzeigen-Annahmestellen in der DDR, für Wirtschaftsanzeigen der VEB Verlag Technik, Oranienburger Str. 13/14, Berlin, 1020, PSF 201. Anzeigenpreisliste Nr. 6; Auslandsanzeigen: Interwerbung GmbH.
Hermann-Duncker-Str. 89, Berlin, DDR, 1157
Erfüllungsort und Gerichtsstand: Berlin-Mitte.

Erscheinungsweise: monatlich 1 Heft

Heftpreis: 4,– M, Abonnementpreis vierteljährlich 12,– M; Auslandspreise sind den Zeitschriftenkatalogen des Außenhandelsbetrieben BUCHEXPORT zu entnehmen

Bezugsmöglichkeiten
DDR
sämtliche Postämter
SVR Albanien
Direktorije Quendrore e Perhapjes dhe Propaganditit te Librit Rruga Konference e Pezes, Tirana
VR Bulgarien
Direkzia R.E.P., 11a Rue Paris, Sofia
VR China
China National Publications Import and Export Corporation, West Europe Department, P. O. Box 88, Beijing
ČSSR
PNS – Ústředni Expedicia a Dovoz Tisku Praha, Slezska 11, 12000 Praha 2
PNS Ústředna Expedicia a Dovoz Tlače, Pošta 002 88547 Bratislava
SFR Jugoslawien
Jugoslovenska Knjiga, Terazije 27, Beograd; Izdavačko Knjižrsko Produzeće MLADOST, Ilica 30, Zagreb
Koreanische DVR
CHULPANMUL Korea Publications
Export & Import Corporation, Pyongyang
Republik Kuba
Empresa de Comercio Exterior de Publicaciones, O'Reily No. 407, Ciudad Habana
VR Polen
C.K.P.iW. Ruch, Towarowa 28, 00-958 Warszawa
SR Rumänien
D.E.P. Bucuresti, Piata Scinteii, Bucuresti
UdSSR
Städtische Abteilungen von Sojuzpečat' oder Postämter und Postkontore
Ungarische VR
P. K. H. I., Külföldi Előfizetési Osztály, P. O. Box 16, 1426 Budapest
SR Vietnam
XUNHASABA, 32, Hai Ba Trung, Hanoi
BRD und Berlin (West)
ESKABE Kommissions-Grossobuchhandlung. Postfach 36, 8222 Ruhpolding/Obb; Helios Literatur-Vertriebs-GmbH, Eichborndamm 141–167, Berlin (West) 52; Kunst und Wissen Erich Bieber OHG, Postfach 46, 7000 Stuttgart 1; Gebrüder Petermann, BUCH + ZEITUNG INTERNATIONAL, Kurfürstenstraße 111 Berlin (West) 30
Österreich
Helios Literatur-Vertriebs GmbH & Co. KG, Industriestraße B/13, A-2345 Brunn am Gebirge
Schweiz
Verlagsauslieferung Wissenschaft der Freihofer AG, Weinbergstr. 109, 8033 Zürich
Alle anderen Länder
örtlicher Buchhandel;
BUCHEXPORT Volkseigener Außenhandelsbetrieb der Deutschen Demokratischen Republik, Leipzig, DDR 7010, Postfach 160; und Leipzig Book Service, Talstraße 29, Leipzig DDR, 7010

INHALT

Bär, H.-P.; Kirpal, A.

30 Jahre integrierte Schaltkreise — Entstehung und Entwicklung der Mikroelektronik. Teil 1

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 122 – 124

Vor 30 Jahren wurde mit der Erfindung des integrierten Schaltkreises (IC) eine entscheidende Entwicklungsstufe der Elektronik eingeleitet. Das Bestreben nach Miniaturisierung und Zusammenfassung elektronischer Schaltungen ist nahezu in jedem Abschnitt der Elektronikentwicklung zu verzeichnen. Mit dem erfolgreichen Bemühen um die weitere Verbesserung der Transistortechnologie in den 50er Jahren, herausragend ist dabei die Ausarbeitung des Si-Planar-Prozesses zu nennen, wurden entscheidende Voraussetzungen für die Entwicklung monolithisch integrierter Halbleiterschaltkreise geschaffen.

Lange, F. H.

Bemerkungen zum Modellbegriff der Systemtechnik

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 125 + 126

Der Modellbegriff ist ein Homonym. Er hat nicht nur in der Systemtheorie und der Systemtechnik unterschiedliche Bedeutung, sondern er wird sogar in gegensetzlichem Sinne verwendet. Auf Aspekte dieses Phänomens geht der vorliegende Beitrag ein.

Müller, K. F.

Impedanz- und Symmetrieverhalten der analogen Teilnehmeranschlußschaltung

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 126 – 128

Der Aufbau der integrierten Schnittstelle zwischen der analogen Teilnehmerleitung und dem Vermittlungsamt wird nach Art der Speisung der Zweidrahtleitung und Art der Gegenkopplung zur Impedanzsynthese unterschieden. Für beide Varianten, die sich in der Praxis durchgesetzt haben, werden Transversal- und Longitudinalimpedanzen sowie Symmetrieverhältnisse angegeben. Es zeigt sich, daß beide untersuchten Varianten gleiche Forderungen an die technologisch kritischen relativen Bauelementetoleranzen stellen.

Trstenský, D.; Beichelt, F.; Sproß, L.

Dynamische Steuerung von Informationsflüssen in Kommunikationsnetzen unter dem Zuverlässigkeits- und Leistungsaspekt

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 129 – 131

Die Leistungsfähigkeit von Kommunikationsnetzen kann wesentlich durch die Optimierung ihrer topologischen Struktur und durch effektive Steuerung des Informationsflusses gesteigert werden. In der Arbeit werden neuere Grundmodelle zur Beherrschung dieser Prozesse vorgestellt.

Sandau, R.; Lorenz, E.

Die quantisierte Fouriertransformation — eine Approximation der DFT

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 131 – 133

Die Relevanz (zyklische Translationsinvarianz) und die Kompaktheit der Leistungsspektren der DFT und der quantisierten Fouriertransformation (QFT) mit 3 bzw. mit 5 Quantisierungsniveaus werden bei Vorliegen einer periodischen Signalform untersucht, wobei sich die Leistungsspektren der QFT als gute Approximation der DFT-Leistungsspektren erweisen. In einem Rechenzeitvergleich zur Durchführung der Transformationen mit Universalprozessoren ergeben sich für die QFT wesentliche Zeitvorteile.

Zühlke, W.

Arhythmische Signalcodierungen und Kompandierung ihrer Zeitbasis

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 134 – 136

Die vorgestellte Zeitbasiskompandierung [9] ermöglicht eine bessere Ausnutzung der Übertragungszeit bei arhythmischer Codierung von Signalen, deren momentanes Spektrum starken Änderungen unterliegt und die Langzeitbandbreite stets nur zu einem geringen Teil ausfüllt. Im Zusammenhang mit der arhythmischen Abtastung ist es ein nichtlineares Verfahren, das neue Wege der Signalbearbeitung zur redundanzarmen Übertragung, Speicherung und Darstellung erschließt.

Hiller, H.; Wobus, C.

Präzisionswinkelmodulation für CPM-Verfahren

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 137 – 140

CPM-Verfahren (Continous Phase Modulation) haben international für die Übertragung von digitalen Signalen an Bedeutung gewonnen. Der Modulationsvorgang kann durch Signalsynthetisierung im Zeitbereich nach verschiedenen Methoden erfolgen. Die bei der Realisierung verschiedener Schaltungen und verschiedener CPM-Verfahren erzielten Ergebnisse werden dargelegt.

Thomanek, H.-J.

Korrelations- und Spektralanalyse digitaler Signale

Nachrichtentech., Elektron., Berlin, 39 (1989) 4, S. 141 – 143

Unter Verwendung des Modells stochastischer Impulsprozesse mit determiniertem Zeitraster gibt der Beitrag einen Überblick zur Berechnung von Autokorrelationsfunktion und Leistungsdichtespektrum zufälliger digitaler Impulssignale. Besondere Bedeutung kommt dabei dem Einfluß des Strukturanteils der Impulsfolge zu. Neben dem trivialen Fall statistisch voneinander unabhängiger Impulse wird die Modellierung der statistischen Bedingungen durch eine Markovkette behandelt.

Freese, K.; Warning, D.

OPTIMOD — Ein universelles Programmsystem zur Unterstützung des Modellierungsprozesses im Schaltkreisentwurf

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 144 – 146

Das dargestellte Programmsystem dient der Parameterextraktion für Netzwerkanalysemodelle integrierter Transistoren. Die Einordnung in den Entwurfsprozeß wird gezeigt und auf die Probleme nichtlinearer Optimierungsverfahren beim Einsatz für praktisch genutzte Transistormodelle wird eingegangen. Am Beispiel des Levenberg-Marquardt Algorithmus ist eine Verfahrenslösung im Detail dargestellt. Auf die numerische Umsetzung wird im Artikel eingegangen. Anschließend sind die Ergebnisse des Verfahrens an 5 verschiedenen Modellierungsproblemen aus der MOS-Technik dargestellt.

Rau, D.

Großsignalverhalten der Avalanchefotodiode beim Einsatz in Breitbandempfängern für Analogsignale

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 147 – 149

Obwohl ursprünglich für digitale Lichtwellenleiter-Nachrichtenübertragungssysteme entwickelt, findet die Avalanchefotodiode (APD) auch in optischen Empfängern für Analogsignale Verwendung. Aufgrund der im Teilnehmeranschlußbereich auftretenden hohen Eingangslichtleistungen können bereits durch die APD nichtlineare Verzerrungen im Empfangssignal verursacht werden. Es werden hierzu ein Lösungsweg zur analytischen Abschätzung der zu erwartenden Klirrdämpfung angegeben sowie mögliche Folgen diskutiert.

Denda, W.

Das Rauschverhalten von Digital-Analog-Wandlern

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 149 – 151

Beim Einsatz von Digital-Analog-Wandlern als programmierbare Referenz-Spannungsquelle stört oft das der Ausgangsspannung überlagerte Rauschen. Die einzelnen Rauschquellen des Wandlers werden selektiert und ihr Beitrag zum Gesamtrauschen definiert. Die gewonnenen Ergebnisse gelten gleichermaßen für die integrale und die spektrale Darstellung des Rauschens. Durch die Wahl der äußeren (nicht integrierten) Baugruppen des Wandlers kann der Anwender die Größe der Ausgangs-Rauschspannung wesentlich beeinflussen. Die dargestellten Meßergebnisse ermöglichen einen Vergleich der Wandler verschiedener Hersteller.

Józwiak, I., Kosmol, U.

Stochastische Modellierung von fehlertoleranten Systemen

Nachrichtentech., Elektron., Berlin 39 (1989) 4, S. 152 + 153

Ausgehend von der programmtechnischen Steuerung von Real-Time-Systemen wird gezeigt, daß bei entsprechender Wahl der Zustandsbeschreibung das Verhalten solcher Systeme durch gesteuerte stochastische Prozesse erfaßbar ist. Als Optimalitätskriterium der Steuerung gilt die erreichte Fehlertoleranz im Sinne der Einhaltung prozeßtypischer Grenzwerte. Der Beitrag ist ein Verfahrensvorschlag, jedoch nicht als vollständig abgeschlossenes Modell zu betrachten.

ISSN 0323–4657 · Nachrichtentech., E
Berlin 39 (1989) 4. S. 121–160